Chapter 1
First Steps

Abstract The basic ideas of MDS are introduced doing MDS by hand. Then, MDS is done using a computer program. The goodness of the MDS configuration is evaluated by correlating its distances with the data.

Keywords MDS configuration · Iteration · Proximities · Dimensional interpretation · Goodness of fit

The basic ideas of MDS are easily explained using a small example. Consider Table 1.1. It contains correlations for the frequencies of different crimes in 50 U.S. states. These correlations show, for example, that if there are many cases of assault in a state, then there are also many cases of murder ($r = 0.81$). In contrast, the murder rate is not correlated with the rate of larceny ($r = 0.06$).

We now scale these correlations via MDS. This means that we try to represent the seven crimes by seven points in a geometric space so that any two points lie the *closer* together the *greater* the correlation of the two crimes that they represent. For this we proceed as follows.

We take seven cards, and write the name of one crime on each of them, from Murder to Auto Theft. These cards are placed on a table in an arbitrary arrangement as shown in Fig. 1.1. We then measure the distances among all cards (Fig. 1.2) and compare these values with the correlations in Table 1.1. This comparison makes clear that the configuration of cards in Fig. 1.1 does not represent the data in the desired sense. For example, the cards Murder and Assault should be relatively close together, because these crimes are correlated with 0.81, whereas the cards Murder and Larceny should be farther apart, as these crimes are correlated with only 0.06. We therefore try to move the cards repeatedly in small steps ("*iteratively*") so that the distances correspond more closely to the data. Figure 1.3 demonstrates in which directions the cards should be shifted, by some small amounts, to improve the correspondence of data and distances.

Since iterative modifications of a given configuration by hand can be fairly tedious and since they do not guarantee that an *optimal* configuration is found in the end,

I. Borg et al., *Applied Multidimensional Scaling*, SpringerBriefs in Statistics,
DOI: 10.1007/978-3-642-31848-1_1, © The Author(s) 2013

Table 1.1 Correlations of crime rates over 50 U.S. states

Crime	Murder	Rape	Robbery	Assault	Burglary	Larceny	Auto theft
Murder	1.00	0.52	0.34	0.81	0.28	0.06	0.11
Rape	0.52	1.00	0.55	0.70	0.68	0.60	0.44
Robbery	0.34	0.55	1.00	0.56	0.62	0.44	0.62
Assault	0.81	0.70	0.56	1.00	0.52	0.32	0.33
Burglary	0.28	0.68	0.62	0.52	1.00	0.80	0.70
Larceny	0.06	0.60	0.44	0.32	0.80	1.00	0.55
Auto theft	0.11	0.44	0.62	0.33	0.70	0.55	1.00

we did not continue these iterations by hand but used an MDS computer program instead. It reports the solution shown in Fig. 1.4.

One such MDS program is PROXSCAL, a module of SPSS. To use PROXSCAL, we first save the correlation matrix of Table 1.1 in a file that we call ‘CorrCrimes.sav’. Then, we only need some clicks in PROXSCAL’s menus or, alternatively, the following commands:

```
GET FILE=‘CorrCrimes.sav’.
PROXSCAL VARIABLES=Murder to AutoTheft
     /PROXIMITIES=SIMILARITIES .
```

The `PROXIMITIES` sub-command informs the program that the data—called *proximities* in this context, a generic term for both similarity and dissimilarity data—must be interpreted as similarities. That is, small data values should be mapped into large distances, and large data values into small distances. No further specifications are needed. The program uses its default specifications to generate an MDS solution. We will show later how these specifications can be changed by the user if desired.

Many other programs exist for MDS. One example with nice graphics is the MDS module in SYSTAT. SYSTAT can be run using commands, or by clicking on various options in a graphical user interface. Having loaded the data file with the correlations, and then calling the MDS procedure, we get the menu in Fig. 1.5. In this menu, we select the variables ‘Murder’, ‘Rape’, etc. and leave all other specifications as they are, except the one for “Regression” (marked by the arrow on the left-hand side), where we request that the MDS program should optimize the relation of data to distances in the sense of a least-squares *linear* regression. (The default is *ordinal* regression which is discussed later; see p. 37f)

Both computer programs—PROXSCAL in SPSS and the MDS module of SYSTAT—generate essentially the same MDS solution for the correlations in Table 1.1. This solution is not only optimal, but also quite good, as Fig. 1.6 shows: The relation of data and distances is almost perfectly linear ($r = -0.99$). Hence, the distances among the points of Fig. 1.3 contain the same information as the correlations of Table 1.1. Expressed differently, the data are properly *visualized* so that one can interpret the distances as empirical evidence: The closer two points in the MDS plane, the higher the correlation of the variables they represent.

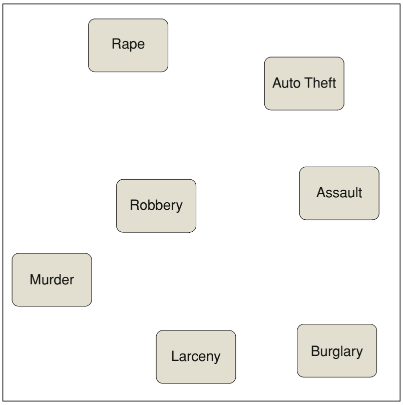


Fig. 1.1 Starting configuration for an MDS of the data in Table 1.1

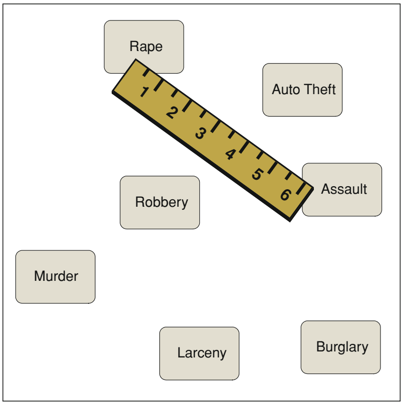


Fig. 1.2 Measuring distances with a ruler

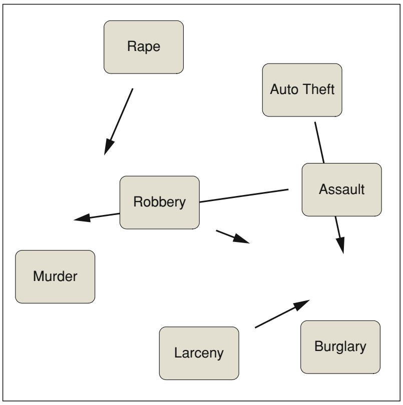


Fig. 1.3 Directions for point movements to improve the MDS configuration

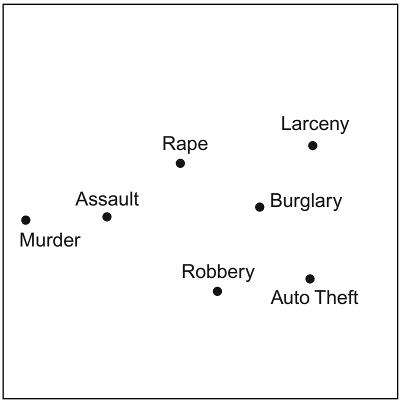
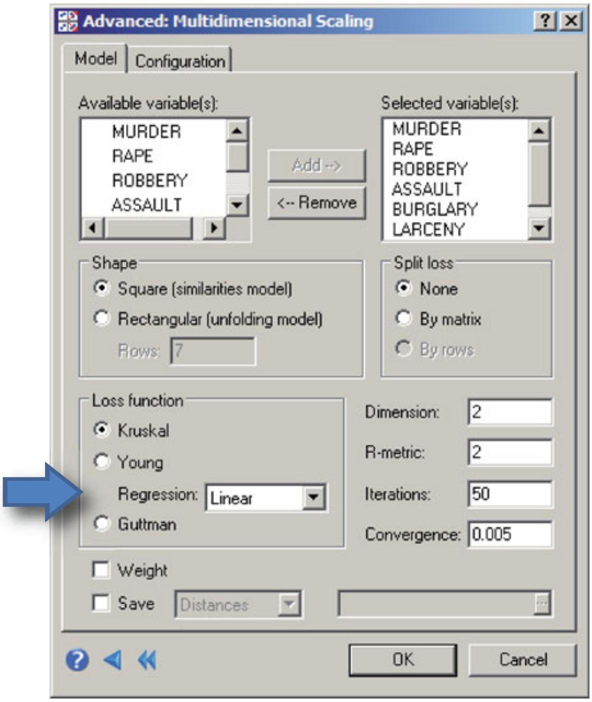


Fig. 1.4 MDS representation of the correlations in Table 1.1 after several iterations

What has been gained by analyzing the crime data via MDS? First, instead of 21 different *numerical* indexes (i.e., correlations), we get a simple *visual* representation of the empirical interrelations. This allows us to actually *see* and, therefore, more easily explore the structure of these data. As shown in Fig. 1.7, the various crimes form certain *neighborhoods* in the MDS plane: Crimes where persons come to harm are one such neighborhood, and property crimes form another neighborhood. This visualizes, for example, that if the murder rate is high in a state, then assault and rape also tend to be relatively frequent. The same applies to property crimes. Robbery lies between these neighborhoods, possibly because violent crimes not only damage the victims' properties but also their bodies.



Advanced: Multidimensional Scaling

Model | Configuration

Available variable(s): MURDER, RAPE, ROBBERY, ASSAULT

Selected variable(s): MURDER, RAPE, ROBBERY, ASSAULT, BURGLARY, LARCENY

Shape
- ● Square (similarities model)
- ○ Rectangular (unfolding model)
- Rows: 7

Split loss
- ● None
- ○ By matrix
- ○ By rows

Loss function
- ● Kruskal
- ○ Young
- Regression: Linear
- ○ Guttman

Dimension: 2

R-metric: 2

Iterations: 50

Convergence: 0.005

□ Weight

□ Save Distances

Fig. 1.5 GUI for the MDS module of SYSTAT

This interpretation builds primarily on the first *principal axis*.[1] This axis corresponds to the horizontal direction of the graph. (Most computer programs for MDS automatically rotate their graphs so that the coordinate axes of MDS plots correspond to principal axes.) The second principal axis is difficult to interpret in this example. On this axis, Larceny and Robbery are farthest apart. Hence, these two crimes might lead us to a meaningful interpretation of the second dimension. Yet, no compelling interpretation seems to offer itself for this dimension: It may simply represent a portion of the "error" component of the data. So, one can ask whether it may suffice to represent the given data in a 1-dimensional MDS space. This is easy to answer: One simply sets "Dimension=1" in the GUI in Fig. 1.5 and then repeats the MDS analysis, leaving all other specifications as before, to get the desired solution.

Figure 1.8 shows the 1-dimensional solution. It closely reproduces the first principal axis of Fig. 1.4. However, its distances correlate with only $r = 0.88$ with the data, i.e. this MDS solution does not represent the data that well. This is also evident from the regression graph in Fig. 1.9, which has a much larger scatter than

[1] The first principal axis is a straight line which runs through the point cloud so that it is closest to the points. That is, the sum the (squared) distances of the points from this line is minimal. Or, expressed differently: The variance of the projections of the points onto this line is maximal. The second major axis is perpendicular to the first and explains the maximum of the remaining variance.

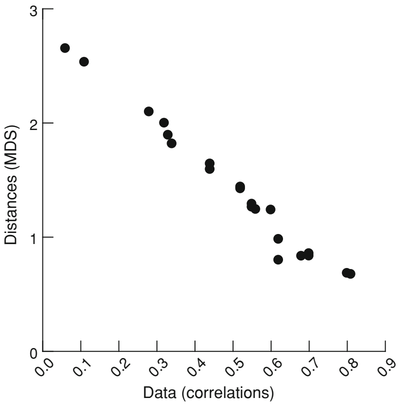


Fig. 1.6 Relation of data in Table 1.1 and distances in Fig. 1.4

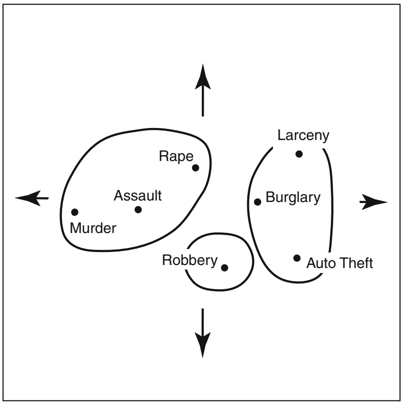


Fig. 1.7 MDS solution with two interpretations: neighborhoods and dimensions

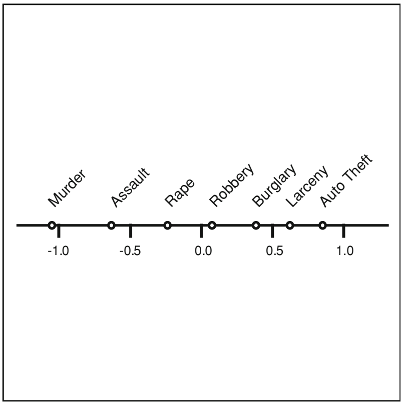


Fig. 1.8 An 1-dimensional MDS solution for the crime data

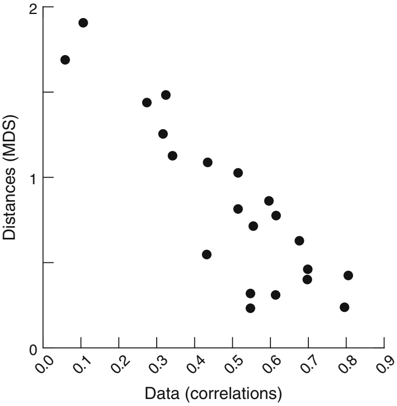


Fig. 1.9 Relation of data in Table 1.1 and distances in Fig. 1.9

the graph for the 2-dimensional MDS solution in Fig. 1.6. One should therefore be cautious when interpreting this configuration, because it is partly misleading. For example, Larceny and Auto Theft correlate much lower ($r = 0.55$) than Larceny and Burglary ($r = 0.80$), but the configuration in Fig. 1.8 does not represent this difference correctly. Rather, the respective two distances are about equal in size.

Summary

Multidimensional scaling (MDS) represents proximity data (i.e., measures of similarity, closeness, relatedness etc.) as distances among points in a multidimensional (typically: 2-dimensional) space. The scaling begins with some starting configuration. Its points are then moved iteratively so that the fit between distances and data is improved until no further improvement seems possible. Computer programs (such as SYSTAT or PROXSCAL) exist for that purpose. The more precisely the data correspond to the distances in the MDS space, the better the MDS point configuration represents the structure of the proximities. If the fit of the MDS solution is good, it can be inspected visually in an attempt to interpret it in terms of content. A popular approach for doing this is to look for dimensions, mostly principal axes, that make sense in terms of what is known or assumed about the objects represented by the points.