Wiley StatsRef: Statistics Reference Online

Multidimensional Unfolding

By *Patrick Mair*[1], *Jan De Leeuw*[2], *and Marcus Wurzer*[3]

Keywords: *fitting distances, multidimensional scaling*

Abstract: Multidimensional unfolding applies distance models and scaling techniques to rectangular matrices of preference and attitude data. We discuss metric and nonmetric variants with optional restrictions on the configurations.

The unfolding model is a geometric model for preference and choice. It locates individuals and alternatives as points in a joint space, and it says that an individual will pick the alternative in the choice set closest to its ideal point. Unfolding originated in the work of Ref. 1 and his students who introduced a unidimensional nonmetric unfolding model. Unfolding models are widely used in **Scaling of Preferential Choice** and **Attitude Scaling**.

The multidimensional unfolding technique computes solutions to the equations of unfolding model. It can be defined as **Multidimensional Scaling** of off-diagonal matrices. This means that the data are dissimilarities between n row objects and m column objects, collected in an $n \times m$ matrix Δ. An important example is preference data, where δ_{ij} indicates, for instance, how much an individual i dislikes object j. In unfolding, we have many of the same distinctions as in general multidimensional scaling: there is unidimensional and multidimensional unfolding and metric and nonmetric unfolding, and there are many possible choices of loss functions that can be minimized.

First, we will look at metric unfolding. We define a multidimensional unfolding loss function and minimize it. In the most basic and classical form, we have the *Stress* loss function

$$\sigma(X, Y) = \sum_{i=1}^{n} \sum_{j=1}^{m} w_{ij}(\delta_{ij} - d_{ij}(X, Y))^2$$

with w_{ij} as optional $n \times m$ weight matrix. This is identical to an ordinary multidimensional scaling problems where the diagonal (row-row and column-column) weights are zero. Or, to put it differently, in unfolding the dissimilarities between different row objects and different column objects are missing. Thus, any multidimensional scaling program that can handle weights and missing data can be used to minimize this loss function. A majorization algorithm can be used to minimize the loss function discussed earlier. Details are in Refs 2 and 3. The flexibility of the majorization algorithm can be used to pose restrictions on the configurations. For instance, we can think of restricting the row (or column) configurations to lie on a sphere by an introducing corresponding spherical projections in each iteration. One can also consider measuring

[1]Harvard University, Cambridge, MA, USA
[2]University of California, Los Angeles, CA, USA
[3]WU Vienna University of Economics and Business, Vienna, Austria

Update based on original article by Jan de Leeuw, Wiley StatsRef: Statistics Reference Online, © 2014, John Wiley & Sons, Ltd

DOI: 10.1002/9781118445112.stat06495.pub2

loss using *Stress*, the sum of squared differences between the squared dissimilarities and squared distances. This has been considered in Ref. 4.

Nonmetric (ordinal) multidimensional unfolding is slightly more complicated. The original techniques proposed by Coombs[1] were purely nonmetric and did not even lead to metric representations. In preference analysis, the prototypical area of application, we often only have ranking information. Each individual ranks a number of candidates, food samples, or investment opportunities. The ranking information is row-conditional, which means we cannot compare the ranks given by individual i to the ranks given by individual k. The order is defined only within rows. Metric data are generally unconditional because we can compare numbers both within and between rows. Because of the paucity of information (only rank order, only row-conditional, and only off-diagonal), the usual Kruskal approach to nonmetric unfolding often leads to degenerate solutions, even after clever renormalization and partitioning of the loss function. In nonmetric unfolding, the *Stress* becomes

$$\sigma(X, Y) = \sum_{i=1}^{n} \sum_{j=1}^{m} w_{ij}(\hat{d}_{ij} - d_{ij}(X, Y))^2$$

with $\hat{d}_{ij} = f(\delta_{ij})$ reflecting a monotone regression on the dissimilarities. Degenerate solutions are characterized by constant d-hats (disparities). Reference 5 identifies constant d-hats using the coefficient of variation and, subsequently, penalize nonmetric transformations of the dissimilarities with small variation. They present a majorization approach for minimizing the adjusted loss function.

One would expect even more problems when the data are not even rank orders but just binary choices. Suppose that n individuals have to choose one alternative from a set of m alternatives. The data can be coded as an *indicator matrix*, which is an $n \times m$ binary matrix with exactly one unit element in each row.

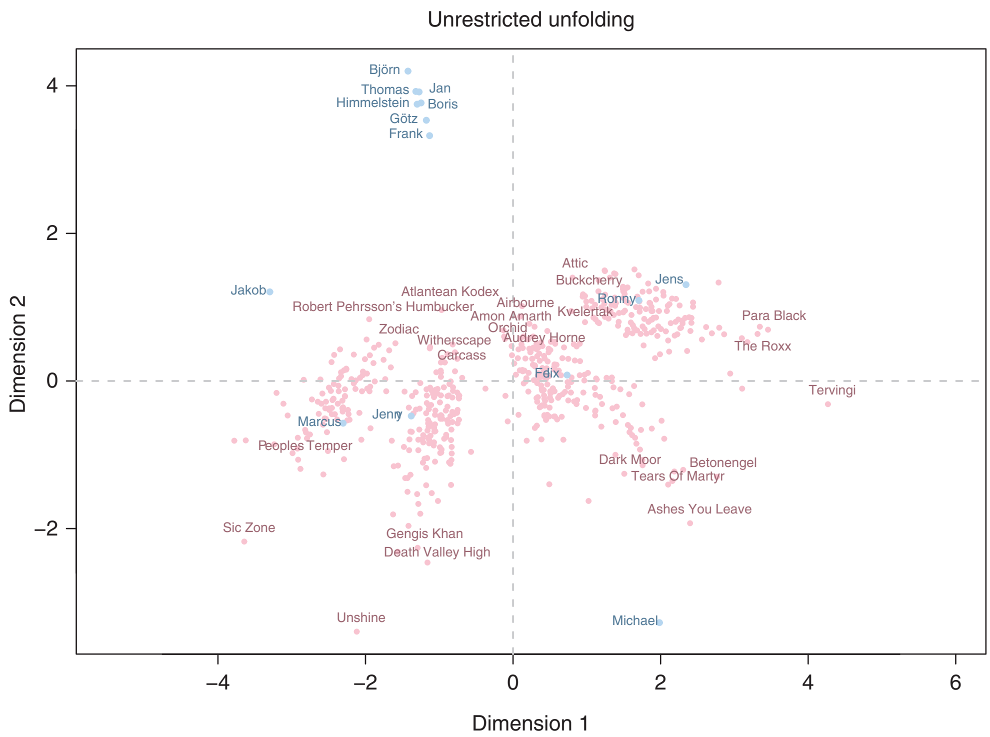

Figure 1. Unrestricted metric unfolding solution.

Wiley StatsRef: Statistics Reference Online, © 2014–2015 John Wiley & Sons, Ltd.
This article is © 2015 John Wiley & Sons, Ltd.
DOI: 10.1002/9781118445112.stat06495.pub2

The unfolding model says there are n points x_i and m points y_j in $\mathbb{R}^p$ such that, if individual i picks alternative j, then $\| x_i - y_j \| \leq \| x_i - y_\ell \|$ for all $\ell = 1, \ldots, m$. The situation becomes more favorable if we have more than one indicator matrix, that is if each individual makes more than one choice. Unfolding in this case can be done by **Correspondence Analysis**.

Now, we present a metric unfolding example based on ratings of records by writers of a music magazine. Each month, writers of the German Heavy Metal magazine *RockHard* rate approximately 50 new records on a scale from 0 (worst) to 10 (best). After collecting the data for the whole year of 2013, we end up with a 14×576 matrix that includes the ratings of 14 writers on 576 bands. A few ratings are missing because not every writer participated in every monthly rating. As a final data preparation step, we subtracted the ratings from 11 such that 1 reflects the highest preference and 11 the lowest preference.

Figure 1 shows the unrestricted two-dimensional unfolding solution with a normalized stress value of 0.23. We only label the top bands of each month and the worst bands of each month. Each "band of the month" is scaled close to the origin, whereas the worst bands of each month are in the periphery of the plot. For instance, the band with the highest rating in 2013 was Atlantean Kodex and the one with the lowest rating Unshine. We see four configuration clusters of bands that roughly correspond to various Heavy Metal subgenres. The closer the reviewers are to each other in the plot, the more similar their ratings, and, therefore, the more similar their musical taste.

As a variant of this example, we fit a restricted two-dimensional unfolding solution on the same data where we pose a circular restriction on the writers (i.e., rows). Naturally, the stress value increases to a value of 0.28. The results are shown in Figure 2.

The image changes. We see two main clusters of bands. Again, the best rated bands are close to the origin and the bands with the lowest ratings are in the periphery. Looking at the distances between writers and bands we see that the bands close to the origin are the ones the writers agree on in terms of ratings.

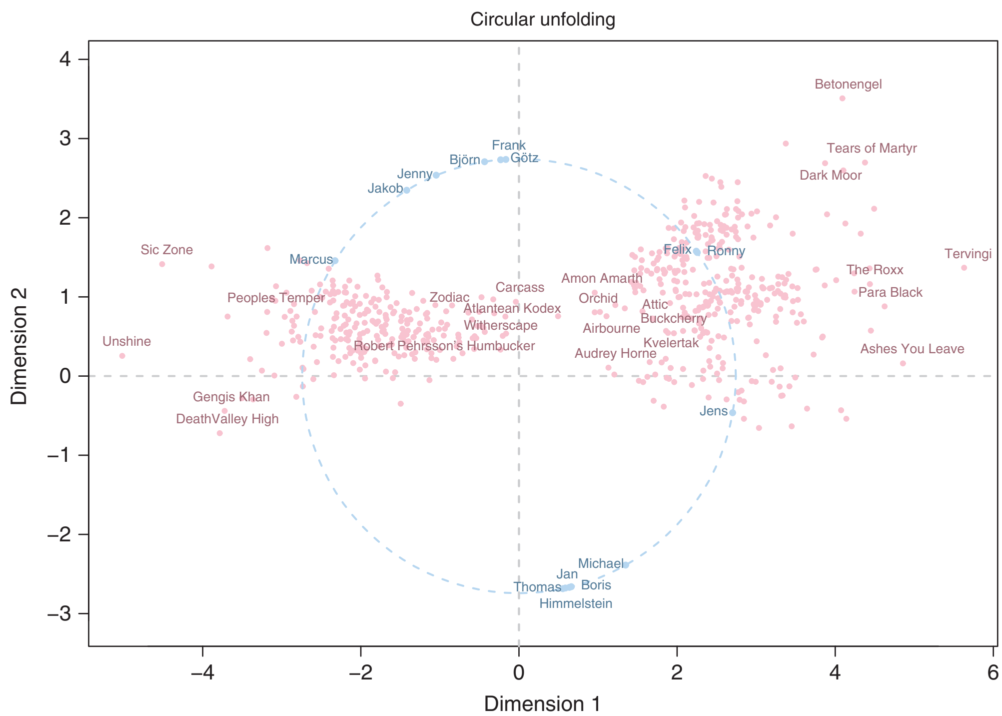

Figure 2. Metric unfolding solution with circular restrictions on the rows.

DOI: 10.1002/9781118445112.stat06495.pub2

In this article, we discussed main approaches to multidimensional unfolding. Some special unfolding models such as ideal point (or external) unfolding, vector model unfolding, and weighted unfolding are discussed in Ref. 3.

References

[1] Coombs, C.H. (1964) *A Theory of Data*, John Wiley & Sons, Inc.
[2] De Leeuw, J. and Mair, P. (2009) Multidimensional scaling using majorization: SMACOF in R. *J. Stat. Softw.*, **31** (3), 1–30.
[3] Borg, I. and Groenen, P.J.F. (2005) *Modern Multidimensional Scaling: Theory and Applications*, 2nd edn, Springer-Verlag.
[4] Takane, Y., Young, F.W., and De Leeuw, J. (1977) Nonmetric individual differences in multidimensional scaling: an alternating least squares method with optimal scaling features. *Psychometrika*, **42**, 7–67.
[5] Busing, F.M.T.A., Groenen, P.J.F., and Heiser, W.J. (2005) Avoiding degeneracy in multidimensional unfolding by penalizing on the coefficient of variation. *Psychometrika*, **70**, 71–98.

DOI: 10.1002/9781118445112.stat06495.pub2