# Smacof at 50: A Manual
# Part 5: Unfolding in Smacof

Jan de Leeuw - University of California Los Angeles

Started December 12 2022, Version of April 29, 2024

**Abstract**

TBD

# Contents



**Note:** This is a working manuscript which will be expanded/updated frequently. All suggestions for improvement are welcome. All Rmd, tex, html, pdf, R, and C files are in the public domain. Attribution will be appreciated, but is not required. The files can be found at https://github.com/deleeuw in the repositories smacofCode, smacofManual, and smacofExamples.

# 1 Introduction

In Multidimensional Unfolding (MDU) the objects of an MDS problem are partitioned into two sets. There is a set of $n$ row-objects and a set of $m$ column-objects, and a corresponding $n \times p$ row-configuration $X$ and $m \times p$ column-configuration $Y$. We minimize stress defined as

$$\sigma(X,Y) := \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(\delta_{ij} - d(x_i,y_j)^2. \tag{1}$$

over both $X$ and $Y$. Here

$$d(x_i,y_j) := \sqrt{(x_i-y_j)'(x_i-y_j)} \tag{2}$$

Thus the within-set dissimilarities are missing, or ignored even if they are available, and only the between-set dissimilarities are fitted by between-set distances.

If we define $Z$ as

$$Z := \begin{array}{c c} & p \\ n & \multirow{2}{*}{$\begin{bmatrix} X \\ Y \end{bmatrix}$} \\ m & \end{array} \tag{3}$$

and $U$ as

$$U := \begin{array}{c c c} & n & m \\ \begin{array}{c} n \\ m \end{array} & \multicolumn{2}{c}{\begin{bmatrix} 0 & W \\ W' & 0 \end{bmatrix}} \end{array} \tag{4}$$

then we can also write

$$\sigma(Z) = \sum_{i=1}^{n+m}\sum_{j=1}^{n+m} u_{ij}(\delta_{ij} - d_{ij}(Z))^2. \tag{5}$$

Data Preferences Type A and Type B Conditional

The unfolding model for preference judgments is often attributed to Clyde H. Coombs (1950), with further developments by Coombs and his co-workers reviewed in C. H. Coombs (1964). After this path-breaking work the digital computer took over, and minimization of loss function (1) and its variations was started by Roskam (1968) and Kruskal and Carroll (1969).

In this manual we are not interested in MDU as a psychological theory, as a model for preference judgments. We merely are interested in mapping off-diagonal dissimilarity relations into low-dimensional Euclidean space, i.e. in making a picture of the data. In some cases (distance completion, distances with errors, spatial basis)

Challenges: degeneracy

# 2 Loss function

## 2.1 Metric

## 2.2 Non-linear

## 2.3 Non-metric

## 2.4 Constraints

# 3 smacofUF

## 3.1 Initial Configuration

$$\sigma(C)=\sigma(\tilde{C}+(C-\tilde{C}))=\sum_{i=1}^{n}\sum_{j=1}^{m}((\delta_{ij}^2-\text{tr }A_{ij}\tilde{C})-\text{tr }A_{ij}(C-\tilde{C}))^2$$

$$\sigma(C)=\sigma(\tilde{C})-2\sum_{i=1}^{n}\sum_{j=1}^{m}(\delta_{ij}^2-\text{tr }A_{ij}\tilde{C})\text{tr }A_{ij}(C-\tilde{C})+\sum_{i=1}^{n}\sum_{j=1}^{m}\{\text{tr }A_{ij}(C-\tilde{C})\}^2$$

From De Leeuw, Groenen, and Pietersz (2006)

$$\sum_{i=1}^{n}\sum_{j=1}^{m}\{\text{tr }A_{ij}(C-\tilde{C})\}^2\leq(n+m+2)\text{tr }(C-\tilde{C})^2$$

Define

$$B(\tilde{C}):=\frac{1}{n+m+2}\sum_{i=1}^{n}\sum_{j=1}^{m}(\delta_{ij}^2-\text{tr }A_{ij}\tilde{C})A_{ij}$$

So we minimize

$$-2\text{ tr }B(\tilde{C})C+\text{tr }C^2-2\text{tr }C\tilde{C}=\text{tr }(C-\{\tilde{C}+B(\tilde{C})\})^2$$

## 3.2 Constraints

### 3.2.1 Centroid Restriction

$$Z=\begin{bmatrix}X\\Y\end{bmatrix}=\begin{bmatrix}I\\D^{-1}G'\end{bmatrix}X=HX$$

Minimize tr $(\overline{Z}-HX)'V(\overline{Z}-HX)$. If there are no further restrictions on $X$ the minimum is attained at $\hat{X}=(H'VH)^{+}H'V\overline{Z}$. Otherwise write $X=\hat{X}+(X-\hat{X})$. Then

$$\text{tr }(\overline{Z}-H\hat{X}-H(X-\hat{X}))'V(\overline{Z}-H\hat{X}-H(X-\hat{X}))=\text{tr }(\overline{Z}-H\hat{X})'V(\overline{Z}-H\hat{X})+\text{tr }(X-\hat{X})'H'VH(X-\hat{X})$$

and we must minimize tr $(X-\hat{X})'H'VH(X-\hat{X})$ for example over $X'H'VHX=I$. That is maximizing $H'VH\hat{X}=H'VHXM$ with $M$ a symmetric matrix of Lagrange multipliers. Thus $M^2=\hat{X}'H'VH\hat{X}$ and $X=\hat{X}(\hat{X}'H'VH\hat{X})^{-\frac{1}{2}}$.

Over tr $X'H'VHX = 1$ we get $H'VH(X - \hat{X}) = \lambda H'VHX$ or $X = (1-\lambda)H'VHX = H'VH\hat{X}$, wich means $X$ is proportional to $\hat{X}$ and we just have to normalize $\hat{X}$ to find $X$.

The constraint $X'V_{11}X = I$ is more difficult to deal with

$$\text{tr}\ (X - \hat{X})'H'VH(X - \hat{X}) =$$

Oblique Procrustus

### 3.2.2 Linearly Restricted Unfolding

#### 3.2.2.1 External Unfolding

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} R & 0 \\ 0 & S \end{bmatrix} \begin{bmatrix} P \\ Q \end{bmatrix}$$

### Normalization Restrictions

$X'V_{11}X = I.$

### 3.2.3 Rank-one Restriction

$Y = za'$

Minimize

$$2\ \text{tr}\ (za' - \overline{Y})'V_{21}(X - \overline{X}) + \text{tr}\ (za' - \overline{Y})'V_{22}(za' - \overline{Y})$$

Removing irrelevant terms

$$2\ z'\{V_{21}(X - \overline{X}) - V_{22}\overline{Y}\}a + a'a.z'V_{22}z$$

Let $H = -V_{21}(X - \overline{X}) - V_{22}\overline{Y}$. Minimize using $z'V_{22}z = 1$.

$$a = H'z = \{\overline{Y}'V_{22} - (X - \overline{X})'V_{12})\}z$$

# 4 Indicator Matrices

Suppose $G_1, \cdots, G_m$ are indicator matrices, all with $n$ rows, but with $k_1, \cdots, k_m$ columns. Their elements are $g^j_{il}$ with $i = 1, \cdots, n$ and $l = 1, \cdots, k_j$. Thus $g^j_{il}$ is either zero or one, and

$$\sum_{l=1}^{k_j} g^j_{il} = 1$$

for all $i = 1, \cdots, n$ and $j = 1, \cdots, m$.

Now define the smacof loss function

$$\sigma(X, Y_1, \cdots, Y_m) = \sum_{j=1}^{m} \sum_{i=1}^{n} \sum_{l=1}^{k_j} w^j_{il}(\delta^j_{il} - d(x_i, y^j_l))^2$$

with $w^j_{il} = g^j_{il}$ and $\delta^j_{il} = 1 - g^j_{il}$. This means we are aiming for $d(x_i, y^j_l) = 0$ whenever $g^j_{il} = 1$ and we do not care what $d(x_i, y^j_l)$ is when $g^j_{il} = 0$.

# 5 Examples

## 5.1 Roskam

## 5.2 Breakfast

## 5.3 Gold

## 5.4 Indicator matrix / Matrices

# References

Coombs, C. H. 1964. *A Theory of Data*. Wiley.

Coombs, Clyde H. 1950. "Psychological Scaling Without a Unit of Measurement." *Psychological Review* 57: 148–58.

De Leeuw, J., P. J. F. Groenen, and R. Pietersz. 2006. "Optimizing Functions of Squared Distances." UCLA Department of Statistics. https://jansweb.netlify.app/publication/deleeuw-groenen-pietersz-u-06/deleeuw-groenen-pietersz-u-06.pdf.

Kruskal, J. B., and J. D. Carroll. 1969. "Geometrical Models and Badness of Fit Functions." In *Multivariate Analysis, Volume II*, edited by P. R. Krishnaiah, 639–71. North Holland Publishing Company.

Roskam, E. E. 1968. "Metric Analysis of Ordinal Data in Psychology." PhD thesis, University of Leiden.