Smacof at 50: A Manual
Part 4: Smacof for Tetrads

Jan de Leeuw - University of California Los Angeles

Started May 19 2024, Version of May 26, 2024

Contents

Note: This is a working paper 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 smacofPC directory of the repositories smacofCode, smacofManual, and smacofExamples.

1 Introduction

The paired comparison of pairs of objects is the simplest and the most basic one of the cartwheel methods for collecting similarity judgments (Coombs (1954)). Pairs of objects from a set $\mathcal{O}$ of n objects are compared, and the data inducate which if the two pairs is the most similar. Torgerson (1958) (p. 261-262) calls this the *method of tetrads*, and we shall adopt this terminology by referring to a pair of pairs of objects $((o_i, o_j), (o_k, o_l))$ as a *tetrad*. This is of course not to be confused with the use of "tetrads" in factor analysis.

In Torgerson's method of tetrads we need replications over occasions (or subjects) to estimate the probability of the $(o_i, o_j) \lesssim (o_k, o_l)$ judgments, which can then turned into scale values using Thurstonianpaired comparison methods. In the smacofPC method we discuss here no replications are needed and we can apply MDS to the data from a single occasion without replications of tetrads.

In the MDS context the response that objects o_i and o_j are more similar than objects o_k and o_l means that we require of our representation X that $d_{ij}(X) \leq d_{kl}(X)$. smacofPC() finds an approximate solution to this highly non-linear and highly non-convex system of inequalities.

smacofPC() is not limited to data collected as comparisons of pairs of objects. Many different data collection methods produce data that are nt of this form, but can be coded as such paired comparisons. Examples are the various versions of the method of triads, the method of conditional rank orders, and the complete or partial ranking of all $\binom{n}{2}$ dissimilarities. In general the comparison data for pairs of pairs do not require transitivity or symmetry and allow for replications of some or all pairs of pairs.

There have been earlier versions of non-metric MDS methods based on comparisons of pairs of objects. Around 1970 various people proposed the *positive orthant method (POM)* (De Leeuw (1968), De Leeuw (1970), Hartmann (1979)), the *absolute value principle (AVP)* (Guttman (1969)), and the *pairwise nonmetric method (PNM)* (Johnson (1973)). Guttman's conference presentation has been republished as Guttman (1979).

The POM, AVP, and PNM methods all use an implicitly normalized least squares loss function with a numerator of the form

$$\lambda_N(X) := \sum \sum w_{ij,kl}\, \sigma_{ij,kl}\, \text{sign}(d_{ij}(X) - d_{kl}(X))|d_{ij}(X) - d_{kl}(X)|^q, \tag{1}$$

where $w_{ij,kl} \geq 0$ and $\sigma_{ij,kl}$ is either minus one, plus one, or zero. In most cases $\sigma_{ij,kl}$ is the signature of the dissimilarities, i.e.

$$\sigma_{ij,kl} = \text{sign}(\delta_{ij} - \delta_{kl}), \tag{2}$$

but that is not necessarily the case, if only because there may not be numerical dissimilarities. It is useful, however, to keep this interpretation in mind. In Johnson's PNM squared distances instead of distances are used, but at the conceptual level that seems just a minor detail.

The denominator of the loss function in all three approaches is

$$\lambda_D(X) := \sum \sum w_{ij,kl}|d_{ij}(X) - d_{kl}(X))|^q, \tag{3}$$

3

and all three algorithms use gradient methods to minimize their loss functions $\lambda(X) = \lambda_N(X)/\lambda_D(X)$.

In both AVP and the more recent version of POM (De Leeuw (2018)) emphasis is on $q = 1$, in which case the loss function is

$$\lambda(X) = \frac{\sum\sum w_{ij,kl}\, \sigma_{ij,kl}(d_{ij}(X) - d_{kl}(X))}{\sum\sum w_{ij,kl}\, |d_{ij}(X) - d_{kl}(X)|}. \tag{4}$$

This can be rewritten in the more compact and computationally more friendly *rearrangement form*

$$\lambda(X) = \frac{\sum r_{ij}d_{ij}(X)}{\max_{r \in \mathcal{R}} \sum r_{ij}d_{ij}(X)}. \tag{5}$$

Here $\mathcal{R}$ is the set of all matrices of the form $r_{ij} = \sum w_{ij,kl}\sigma_{ij,kl}$, with $\sigma_{ij,kl}$ varying over all possible signatures (or, equivalently, over the convex set of all matrices with entries between minus and plus one).

As we shall see, the loss function used in smacofPO is an explicitly normalized version of equation (1) with $q = 2$. The explicit normalization is a major difference with POM/AVP/PNM, which use implicit normalization of stress with denominator (3). The loss function used for our initial estimate is an explicitly normalized version of equation (1) that has $q = 1$ and uses squared distances.

The MAXSCAL method proposed by Takane (1978b), Takane (1978a) is very different from the POM, AVP, PNM trio, and consequently also different from smacofPC. The loss function is not least squares, at least not conceptually, but is based on the likelihood function for a simple choice model for the tetrad responses. Computationally the negative likelihood is minimized by using a Gauss-Newton approximation to the Hessian, which leads in the usual way to iterative reweighted least squares.

Another recent excursion into scaling tetrads is Agarwal et al. (2007). They do not use a least squares loss function and rely on semidefinite programming methods to find an approximate solution to the system of inequalities.

4

2 Loss Function

To introduce the smacofPO loss function we start with a general non-metric MDS problem with loss function

$$\sigma(X, \hat{D}_1, \cdots, \hat{D}_s) := \sum_{r=1}^{R} \sigma_r(X, \hat{D}_r), \tag{6}$$

with

$$\sigma_r(X, \hat{D}_r) := \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ijr}(\hat{d}_{ijr} - d_{ij}(X))^2. \tag{7}$$

As usual, the symbol := is used for definitions. Our loss function is similar to the one used by Roskam (1970) for the method of triads (article republished in Roskam (1979)).

The *weights* $W_r = \{w_{ijr}\}$ are known non-negative numbers and $D(X) := \{d_{ij}(X)\}$ is a matrix of Euclidean distances between the rows of the $n \times p$ *configuration* $X = \{x_{is}\}$, which are interpreted as n points in $\mathbb{R}^p$.

Loss function (6) must be minimized over *configurations* X and over the R matrices of *disparities* $\hat{D}_r = \{\hat{d}_{ijr}\}$. The minimization problem has some constraints, both on X and on the $\hat{D}_r$. We require that $X \in \mathcal{X} \subseteq \mathbb{R}^{n \times p}$. Here $\mathcal{X}$ is the set of *column-centered* (columns add up to zero) $n \times p$ matrices that are *normalized* by

$$\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}^{\star} d_{ij}^2(X)) = 1, \tag{8}$$

where

$$w_{ij}^{\star} := \sum_{r=1}^{R} w_{ijr}. \tag{9}$$

Roskam (1970) uses implicit normalization in his method of triads (and the MINITRI program) by defining loss as the ratio of (7) and the sum of squares from (8). By Kruskal and Carroll (1969) and De Leeuw (1975) this gives the same solution, up to a scale factor, as our explicit normalization.

The *disparities* $\hat{D}_r = \{\hat{d}_{ijr}\}$ are required to satisfy $\hat{D}_r \in \mathcal{C}_r$. The $\mathcal{C}_r$ are polyhedral convex cones, which are subcones of the cone of non-negative matrices. Each of the cones is defined by partial orders $\leq_r$ over the elements of the $\hat{D}_r$. In general, neither the (known) weight matrices W_r nor the (unknown) disparity matrices $\hat{D}_r$ need to be symmetric and/or hollow (i.e. have zero diagonal).

The *data* of the MDS problem are the weights W_r and the cones $\mathcal{C}_r$. Each pair $(W_r, \mathcal{C}_r)$ is called a *slice* of the data. Comparisons of pairs are only within slices. In the MDS literature this is also known as conditional rank orders. The *unknowns* or *parameters* of the problem are $X \in \mathcal{X}$ and the $\hat{D}_r \in \mathcal{C}_r$.

To minimize loss over $X \in \mathcal{X}$ for the current best value of the $\hat{D}_r$. This subproblem is simplified by using the least squares partitioning

$$\sigma(X, \hat{D}_1, \cdots, \hat{D}_s) = \sum_{r=1}^{R} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ijr}(\hat{d}_{ijr} - \hat{d}_{ij}^{\star})^2 + \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}^{\star}(\hat{d}_{ij}^{\star} - d_{ij}(X))^2, \tag{10}$$

5

where

$$\hat{d}_{ij}^{\star} = \frac{\sum_{r=1}^{R} w_{ijr}\hat{d}_{ijr}}{\sum_{r=1}^{R} w_{ijr}}. \tag{11}$$

2.1 Slice-Independence

We will assume throughout that $w_{ijr} = w_{ij}\epsilon_{ijr}$, where $\epsilon_{ijr} = 1$ is either zero or one. If ϵ_{ijr} is one, we say that pair (i, j) *participates* in slice r. Thus $\epsilon_{ijr} = 0$ for all pairs that do not particpate. We refer to the assumption $w_{ijr} = w_{ij}\epsilon_{ijr}$ on the weights as the *slice-independent* case.

To see the consequences of slice-independence for our equations we define

$$\mathcal{J}_r := \{(i, j) \mid \epsilon_{ijr} = 1\} \tag{12}$$

so that

$$\sigma_r(X, \hat{D}_1, \cdots, \hat{D}_s) = \sum_{r=1}^{s} \sum_{(i,j)\in\mathcal{J}_r} w_{ij}(\hat{d}_{ijr} - d_{ij}(X))^2 \tag{13}$$

Equation (9) gives $w_{ij}^{\star} = w_{ij}\epsilon_{ij}^{\star}$, where $\epsilon_{ij}^{\star}$ is the number of times pair (i, j) occurs in the R slices. A set of slices is *balanced* if all $\epsilon_{ij}^{\star}$ are equal. Also, from (11),

$$\hat{d}_{ij}^{\star} = \frac{\sum_{r=1}^{R} \epsilon_{ijr}\hat{d}_{ijr}}{\sum_{r=1}^{R} \epsilon_{ijr}} \tag{14}$$

which does not depend on the w_{ij}.

It follows that in our computations we have to deal with various different sets of weights. There are the w_{ijr}, the $w_{ij}^{\star}$, the w_{ij}, and the $\epsilon_{ij}^{\star}$. In the first ALS subproblem where we minimize over the $\hat{D}_r$ for fixed X we use equation (13), i.e. we use the w_{ij}. If we minimize over X for fixed $\hat{D}_r$ we use (10), which means we use $w_{ij}^{\star} = w_{ij}\epsilon_{ij}^{\star}$ and $\hat{d}_{ij}^{\star}$ given by (14). Of course this all simplifies if $w_{ij} = 1$ for all pairs (i, j) (the *unweighted* case).

2.2 Initial Configuration

In metric and non-linear MDS the default initial configuration is the classical Torgerson metric MDS solution. That is not available for smacofPC, because there are no numerical dissimilarities. In De Leeuw (1970) (section 5.1) and De Leeuw (1973) (republished as De Leeuw (1984)), chapter 4, a different eigenvalue-eigenvector based initial solution is proposed. We discuss a somewhat modernized version here. It is sometime known as the *maximum sum method*. The data are an order over a number of tetrads.

$$\omega(X) := \sum \sum_{(i,j)\prec(k,l)} (d_{ij}^2(X) - d_{kl}^2(X)), \tag{15}$$

where $(i, j) \prec (k, l)$ means that we want $d_{ij}(X) \leq d_{kl}(X)$. We want all terms in $\omega(X)$ to be positive, but for purposes of the initial configuration we relax this to wanting their sum to be large. This explains the "maximum sum" name.

The sum (15) can be written as the quadratic form $\omega(X) = \text{tr } X'A^\star X$, with

$$A^\star := \left\{ \sum_{(i,j)} \sum_{(k,l)} (A_{ij} - A_{kl}) \right\} \tag{16}$$

Note that $A^\star$ is symmetric and doubly-centered. Moreover its trace is zero, and consequently it has one zero, some negative, and some positive eigenvalues.

Because X is column centered we have $\omega(X) = \text{tr } X'(A^\star + \theta J)X$ where J is the centering matrix $I - n^{-1}ee'$ and θ is arbitrary. For the non-zero eigenvalues we have $\lambda_s(A^\star + \theta J)) = \lambda_s(A^\star) + \theta$, and thus for $\theta \geq -\lambda_{\min}(A^\star)$ the matrix $A^\star + \theta J$ is positive semi-definite.

Of course maximizing ω over all X does not make sense, because by making X larger we make ω larger. Thus the supremum over all X is $+\infty$ and the maximum is not attained. We need some kind of normalization. The obvious choice is $\text{tr } X'X = 1$, but unfortunately that does not work. It gives a rank-one solution with all columns of X equal to the eigenvector corresponding with the dominant eigenvalue of $A^\star$. Instead we use $\text{tr } (X'X)^2 = 1$. This gives the solution $X = K\Lambda^{\frac{1}{2}}$ with Λ the largest p eigenvalues of $A^\star$ (assumed to be non-negative) and K the corresponding normalized eigenvectors. This is our version of the Torgerson initial solution for non-metric MDS.

We have been deliberately vague about what to do if the number of positive eigenvalues of $A^\star$ is less than p, which is of course a problem the Torgerson metric MDS solution has as well. In the program we simply choose θ equal to $-\lambda_p(A^\star)$ if $\lambda_p(A^\star) < 0$. We expect the problem to be rare, and the actual choice of θ to be fairly inconsequential.

Maximizing the maximum sum loss function is equivalent to minimizing the loss function $\lambda_N(X)$ in (1) with $w_{ij,kl} = 1$ and $q = 1$ (and with squared distances instead of distances). The maximum sum method was proposed, together with POM, in De Leeuw (1968). The inspiration came mostly from Guttman (1946), so it is no surprise that Guttman (1969) proposes essentially the same initial configuration for his AVP technique.

The method for the special case in which there are numerical (dis)similarities has been around even longer. It is discussed in the one-dimensional case as the "fourth method of quantification" in Hayashi (1952). Suppose Δ is a symmetric non-negative matrix of dissimilarities. Then, in the multidimensional extension proposed in Okamoto and Isogai (1878), we maximize

$$\sum \sum \delta_{ij}^2 d_{ij}^2 = \text{tr } X'BX, \tag{17}$$

with explicit normalization of X. Here

$$B := \sum \sum \delta_{ij}^2 A_{ij}. \tag{18}$$

Thus B is positive semi-definite, and if Δ is irreducible it is of rank $n - 1$. Okamoto and Isogai (1878) observe in their Theorem 3, as we have above. that requiring $\text{tr } X'X = 1$ generally leads to a one dimensional solution. Only if the largest eigenvalue of B has multiplicity $r > 1$ an r-dimensional solution is possible. We removed this problem by using the explicit normalization $\text{tr } (X'X)^2 = 1$.

We should emphasize that in Hayashi's method similarities in stead of dissimilarities are used, and that the (dis)similarities are not squared. It seems natural to use squares dissimilarities in .., also because Okamoto and Isogai (1878) refer to earlier results (in Japanese), where Hayashi's fourth method compares unfavorably with the Torgerson classical MDS solution which is of course based on squared dissimilarities. Also note again that in the tetrad case we do not need a numerical measure of similarity or dissimilarity.

2.3   Pairwise Monotone Regression

Suppose datum r says that that $(i, j) \prec (k, l)$. In the slice independent case $w_{ijr} = w_{ij}$ and $w_{klr} = w_{kl}$ can be non-zero and all other elements of W_r are zero.

If $(i, j) \prec (k, l)$

$$\sigma_r(X, \hat{D}_r) = w_{ij}(\hat{d}_{ijr} - d_{ij}(X))^2 + w_{kl}(\hat{d}_{klr} - d_{kl}(X))^2 \tag{19}$$

Loss (19) must be minimized over $\hat{d}_{ijr} \leq \hat{d}_{klr}$. If $d_{ij}(X) \leq d_{kl}(X)$ then $\hat{d}_{ijr} = d_{ij}(X)$ and $\hat{d}_{klr} = d_{kl}(X)$. Otherwise

$$\hat{d}_{ijr} = \hat{d}_{klr} = \frac{w_{ij}d_{ij}(X) + w_{kl}d_{kl}(X)}{w_{ij} + w_{kl}}. \tag{20}$$

Thus loss component (19) is zero if the order of $d_{ij}(X)$ and $d_{kl}(X)$ is the same as the order dictated by the data and

$$\frac{w_{ij}w_{kl}}{w_{ij} + w_{kl}}(d_{ij}(X) - d_{kl}(X))^2 \tag{21}$$

So far we have only considered the forced-choice situation in which the data is a choice of one of the two pairs. If we allow for the alternative that (i, j) and (k, l) are equally similar then we can choose between two different approaches. In the *primary approach* we incur no loss for this pair, no matter what $d_{ij}(X)$ and $d_{kl}(X)$ are. In the *secondary approach* we require that $\hat{d}_{ijr} = \hat{d}_{klr}$ and consequently we use equation (20) and add to the loss if $d_{ij}X) \neq d_{kl}(X)$.

From equation (21) we see that the minimum of the loss function over $\hat{D}$ for given X is the same as the numerator of Johnson's PM loss function (except that for PM all weights are one).

3 Program

3.1 Parameters

The smacofPC function in R has the following parameters (with default values).

```r
smacofPC <- function(data,
                     nobj = max(data),
                     ndim = 2,
                     wmat = NULL,
                     xold = NULL,
                     labels = NULL,
                     width = 15,
                     precision = 10,
                     itmax = 1000,
                     eps = 1e-10,
                     verbose = TRUE,
                     kitmax = 5,
                     keps = 1e-10,
                     kverbose = 0,
                     init = 1,
                     ties = 0)
```

- If xold is non-null then an initial configuration matrix must be provided.
- If labels is non-null then a character vector of plot labels must be provided.
- width and precision are relevant for the format of (optional) major iteration output.
- itmax and eps determine when the major iterations stop.
- If verbose = TRUE itel and stress for each major iteration are written to stdout.
- kitmax and keps determine the number of inner Guttman transform iterations between two monotone regressions.
- If kverbose = TRUE then itel and stress for each inner iteration are written to stdout.
- If init = 1 the maximum sum initial configuration is computed, if init = 2 a random initial configuration is used.
- Ties is either 0, 1, or 2. If ties = 0 the data are forced choice, no ties are allowed. If ties = 1 or ties = 2 the the primary or secondary approach to ties is used.

3.2 Input

The data are either a four column (if ties = 0) or a five column (if ties = 1 or ties = 2) matrix (or data frame). Here is an example of a data matrix with ties = 2 from the hoogeveen example (see below). The first four columns are indices. Row one, for example, tells us that $\delta_{23} \leq \delta_{34}$. Row two has an entry in the fifth column and tells us that $\delta_{23} = \delta_{45}$, and that we are supposed to use the second approach to ties, i.e. require $\hat{d}_{23} = \hat{d}_{45}$.

```
##     V1 V2 V3 V4 V5
## 1    2  3  3  4  0
```

```
## 2     2   3   4   5   2
## 3     2   3   3   5   0
## 4     2   3   2   4   0
## 5     4   5   1   2   0
## 6     2   5   1   2   2
## 7     3   5   4   5   0
## 8     1   3   1   2   0
## 9     1   2   2   4   0
## 10    2   5   2   4   0
## 11    3   5   3   4   0
## 12    1   3   2   3   0
## 13    1   2   2   5   0
## 14    3   5   2   5   0
## 15    1   5   1   4   0
## 16    1   3   3   4   0
## 17    3   5   1   2   0
## 18    1   5   1   4   0
## 19    3   5   2   3   2
## 20    2   5   2   4   0
```

- The data can have replications of some or all comparisons.
- The data are not necessarily consistent with any partial order, i.e. there can be intransivities and asymmetries.
- It is possible to use the secondary approach to ties for some comparisons and the primary approach for others.
- For more information on data generation see section 4 on utilities.

3.3   Algorithm

The smacofPO algorithm is standard alternating least squares, with one or more Guttman transforms alternated with pairwise monotone regression.

3.4   Output

```r
h <- list(
  nobj = nobj,
  ndim = ndim,
  snew = snew,
  itel = itel,
  xnew = xnew,
  dhat = dhat,
  dmat = dmat,
  wmat = wmat,
  esum = esum,
  wsum = wsum,
```

```
    labels = labels
)
```

The member that may not be obvious are esum and wsum. esum is a matrix with the $\epsilon_{ij}^{\star}$, and wsum is the elementwise product of esum and wmat.

4 Utilities

4.1 Data Generation and Collection

Let us first address the elephant in the room. Even for moderate n there are a lot of tetrads, and in behavioural science it quickly becomes impossible to present all of them to a subject, even if that subject is paid or is an undergraduate psychology student.

If we only consider distinct pairings of distinct pairs there are already $\binom{\binom{n}{2}}{2}$ tetrads, which is of the order $\frac{1}{4}n^4$. Here is a little table.

```
##    3          3
##    4         15
##    5         45
##    6        105
##    7        210
##    8        378
##    9        630
##   10        990
##   11       1485
##   12       2145
##   13       3003
##   14       4095
##   15       5460
##   16       7140
##   17       9180
##   18      11628
##   19      14535
##   20      17955
```

There are a number of ways to deal with this fundamental problem. If we take the psychophysical point of view that subjects are merely replications then we can use multiple subjects, each handling a subset of the tetrads. Alternatively, we can select a random subset of the total set of tetrads. Or we could use a design to select a preferably balanced subset. If the data are not behavioural and machine generated a complete set of tetrads may be feasible.

Two utility functions are included in the package to collect data and organize them into the appropriate format. They are smacofMakeAllPairs <- function(names, ties = 0) and smacofMakeRandomPairs <- function(names, nrandom, ties = 0).

Since 1897 Dutch children were taught to read using a leesplankje (reading board), named after its originator, primary school teacher M.B. Hoogeveen. It was still in use when I went to primary school in the nineteen fifties. It probably has been replaced by more modern tools by now, and leesplankjes are now collectors items for boomers. A prototypical example is in figure 1.

For our stimulus set we selected the five humans Wim, Zus, Jet, Teun, Gijs. Starting smacof-MakeAllPairs() generates a sequence of tetrads on stdout, giving the subject an opportunity to

Figure 1: Hoogeveen's Leesplankje.

choose the most similar pair (or to decide that both pairs are equally similar). On stdout we see (with responses)

```
> smacofMakeAllPairs(names)
(Gijs,Teun) and (Gijs,Wim)
most similar pair: 2
(Teun,Wim) and (Gijs,Jet)
most similar pair: 2
(Gijs,Zus) and (Jet,Zus)
most similar pair: 2
(Wim,Jet) and (Gijs,Wim)
most similar pair: 1
(Teun,Zus) and (Teun,Jet)
most similar pair:
```

At the same time the same pairs are show on the graphics device, in the form shown in figure 2.

Figure 2: A Tetrad

This continues until all tetrads are shown (in the case of smacofMakeAllPairs()) or a given number of random pairs (without replacement) has been shown (in the case of smacofMakeRandomPairs()).

As we discussed in the introduction smacofPC() is not limited to data collected as tetrads. After suitable manipulation it can also handle data collected by the method of triads or by conditional or complete rank orders. In particular if we have a matrix of dissimilarities between n objects we can convert it to a large number of comparisons of tetrads. This is done by the function smacofMakePairsFromDelta <- function(delta, ties = 1).

Alternatively we can code the order of dissimilarities using only $\binom{n}{2} - 1$ tetrads, by using only those pairs that are consecutive in the ordering of pairs. This is done by smacofOrderPairsFromDelta <- function(delta, ties = 1).

4.2 Plots

smacofPC does not have a Shepard plot, because there are no dissimilarities to put on the horizontal axes. We do have

```
smacofConfigurationPlot <-
  function(h,
           main = "ConfigurationPlot",
           dim1 = 1,
           dim2 = 2,
           pch = 16,
           col = "RED",
           cex = 1.0)
```

and

```
smacofDistDhatPlot <- function(h,
                               fitlines = TRUE,
                               colline = "RED",
                               colpoint = "BLUE",
                               main = "Dist-Dhat Plot",
                               cex = 1.0,
                               lwd = 2,
                               pch = 16)
```

5 Examples

5.1 Hoogeveen

A single subject (me) ran smacofMakeAllPairs() on the five leesplankje names. This produced 45 tetrads.

Convergence is to stress 3.1647986×10^{-9} in 296 iterations. The curved solution Teun-Gijs-Wim-Jet-Zus goes from old-male to young-female. The 45 inequalities in the data can be satisfied exactly.

INSERT FIGURE 3 ABOUT HERE

INSERT FIGURE 4 ABOUT HERE

5.2 Parties

I left The Netherlands in 1987 and have not followed their local political news. So I am not really an expert on current Dutch politics. Nevertheless I was the only subject generating the data (50 random tetrads) for this example. The ten parties I used are

- SP Socialists
- GL Greens
- PvdD Party for the Animals
- PvdA Labour
- D'66 Neither fish nor fowl
- CDA Christian Democrats
- VVD European-style Liberals
- CU More Christian Democrats
- SGP Right-wing Christians
- FVD Right-wing Populists

For more details, see this Wikipedia article.

Convergence is to stress 1.1150739×10^{-9} in 135 iterations. The solution is in an almost equilateral triangle with the three extremes (FVD, SP, SGP) in the vertices. The interior of the triangle shows a social democratic (PvdA, PvdD, GL), a liberal (VVD, D'66), and a christian democrat (ARP, CU, CDA) cluster, roughly ordered from left-wing on the right of the plot to right-wing on the left of the plot. Again, all 50 inequalities can be fitted exactly.

We can check how sparse and balanced the randomly generated "design" is by printing the $\epsilon_{ij}^{\star}$.

```
##        [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10]
## [1,]     0    2    3    3    2    4    2    1    0     2
## [2,]     2    0    2    3    1    2    1    4    3     1
## [3,]     3    2    0    3    3    3    4    2    1     0
```

```
##  [4,]    3   3   3   0   0   3   1   4   4   2
##  [5,]    2   1   3   0   0   2   0   3   2   6
##  [6,]    4   2   3   3   2   0   2   0   4   5
##  [7,]    2   1   4   1   0   2   0   1   2   1
##  [8,]    1   4   2   4   3   0   1   0   2   2
##  [9,]    0   3   1   4   2   4   2   2   0   2
## [10,]    2   1   0   2   6   5   1   2   2   0
```

with row sums 19, 19, 21, 23, 19, 25, 14, 19, 20, 21. For an all-pairs-of-pairs design the elements of this matrix would have been 44. So the random "design" is very sparse but decently balanced.

INSERT FIGURE 5 ABOUT HERE

INSERT FIGURE 6 ABOUT HERE

5.3   Ekman (1954)

The ekman data are dissimilarities between 14 colors. This means they can be expanded into 4095 tetrads, which can then be analyzed by smacofPC(). We do not necessarily recommend this practice, because the pairwise coding involves a great deal of dependence and redundancy. The smacofPC() program is really intended for sequences of independent tetrads judgments. Nevertheless we will use these expanded ekman dissimilarities as an example.

Convergence is to stress 0.0087145 in 82 iterations. The solution is the familiar color circle, except that, for some reason that I do not understand yet, the circle is crumpled at the high frequency end.

INSERT FIGURE 7 ABOUT HERE

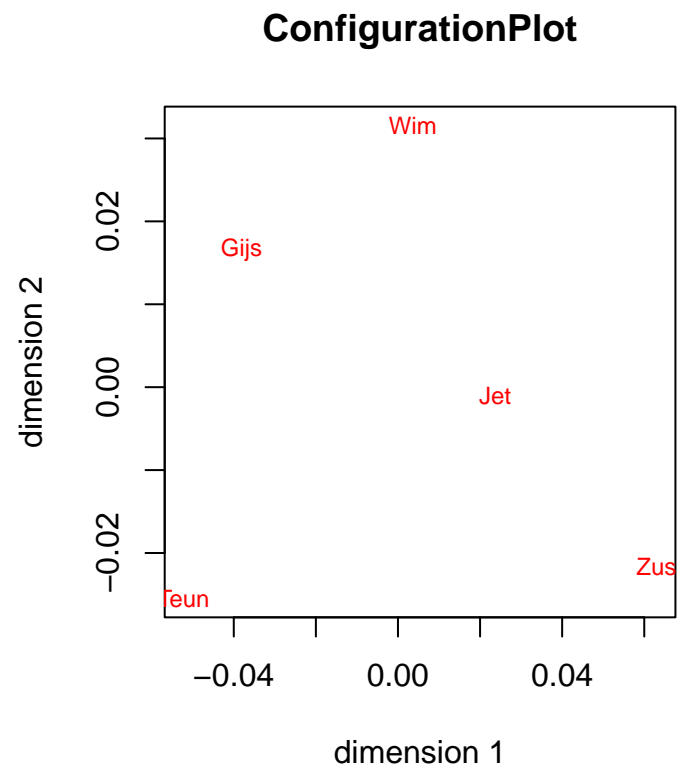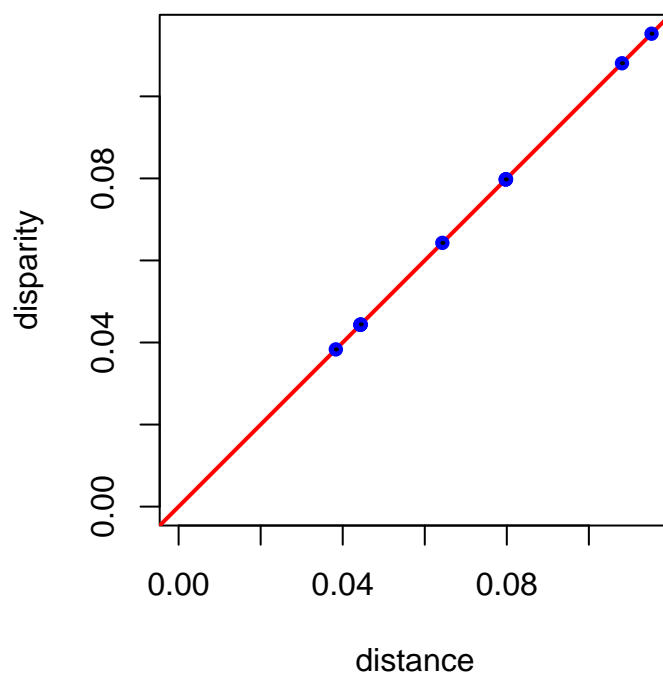INSERT FIGURE 8 ABOUT HERE

6 Figures

ConfigurationPlot

Figure 3: Leesplankje

Dist−Dhat Plot

Figure 4: Leesplankje

ConfigurationPlot
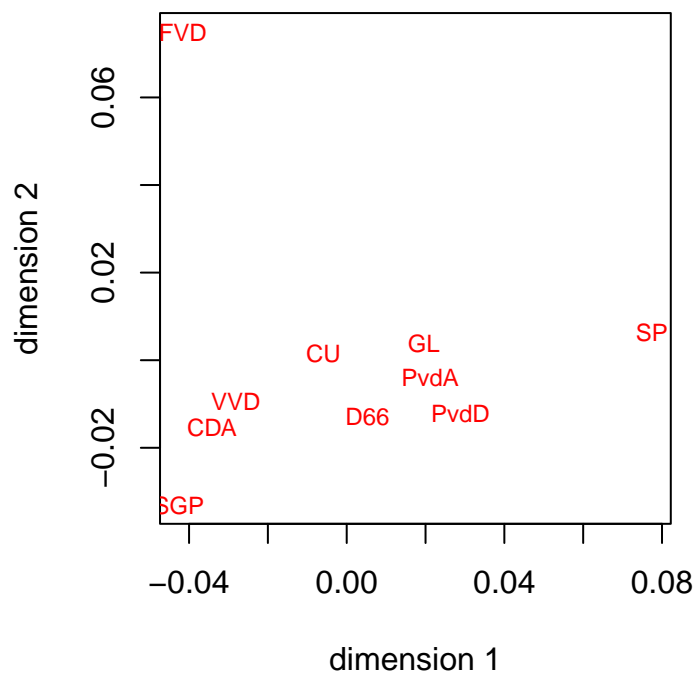
Figure 5: Dutch political parties

Dist−Dhat Plot
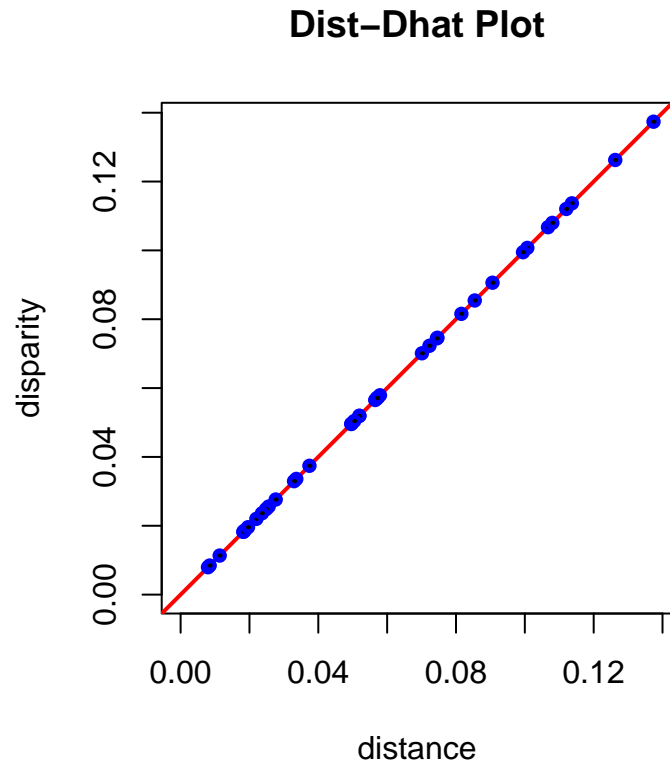
Figure 6: Dutch political parties
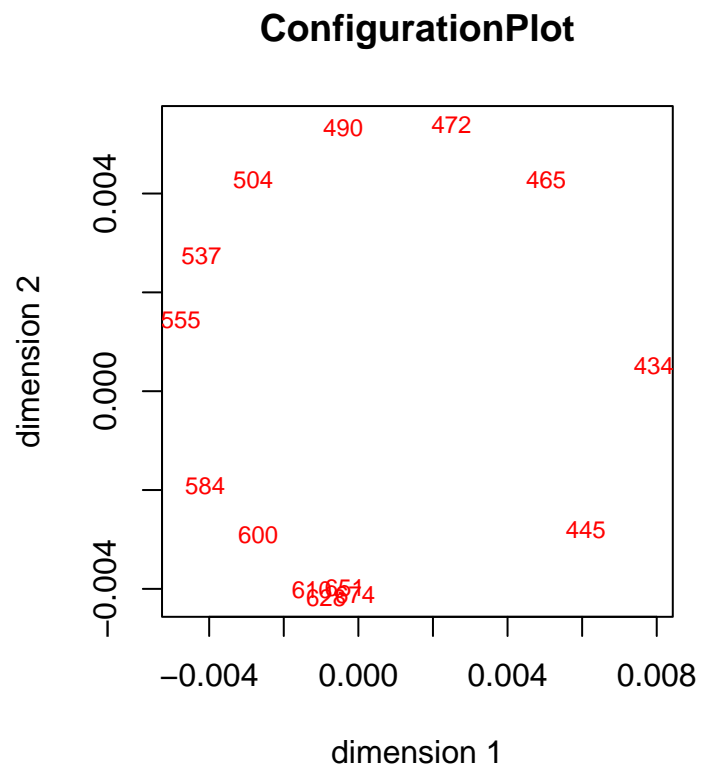
ConfigurationPlot

Figure 7: Ekman color data
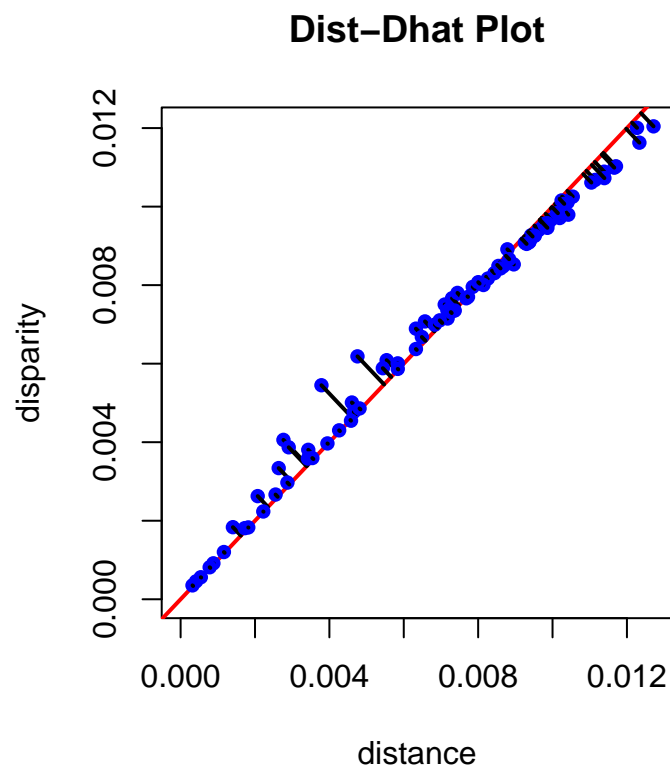
Dist−Dhat Plot

Figure 8: Ekman color data

References

Agarwal, S., J. Wills, L. Cayton, G. Lanckriet, D. Kriegman, and S. Belongie. 2007. "Generalized Non-Metric Multidimensional Scaling." In *Proceedings of the Eleventh International Conference on Artificial Intelligence and Statistics*, 11–18.

Coombs, C. H. 1954. "A Method for the Study of Interstimulus Similarity." *Psychometrika* 19: 183–94.

De Leeuw, J. 1968. "Nonmetric Multidimensional Scaling." Research Note 010-68. Department of Data Theory FSW/RUL. https://jansweb.netlify.app/publication/deleeuw-r-68-9/deleeuw-r-68-g.pdf.

———. 1970. "The Positive Orthant Method for Nonmetric Multidimensional Scaling." Research Report 001-70. Leiden, The Netherlands: Department of Data Theory FSW/RUL. https://jansweb.netlify.app/publication/deleeuw-r-70-a/deleeuw-r-70-a.pdf.

———. 1973. "Canonical Analysis of Categorical Data." PhD thesis, University of Leiden, The Netherlands. https://jansweb.netlify.app/publication/deleeuw-b-73/deleeuw-b-73.pdf.

———. 1975. "A Normalized Cone Regression Approach to Alternating Least Squares Algorithms." Department of Data Theory FSW/RUL.

———. 1984. *Canonical Analysis of Categorical Data*. Leiden, The Netherlands: DSWO Press. https://jansweb.netlify.app/publication/deleeuw-b-84/deleeuw-b-84.pdf.

———. 2018. "The Positive Orthant Method ." 2018. https://jansweb.netlify.app/publication/deleeuw-e-18-d/deleeuw-e-18-d.pdf.

Ekman, G. 1954. "Dimensions of Color Vision." *Journal of Psychology* 38: 467–74.

Guttman, L. 1946. "An Approach for Quantifying Paired Comparisons and Rank Order." *Annals of Mathematical Statistics* 17: 144–63.

———. 1969. "Smallest Space Analysis by the Absolute Value Principle." In *Proceedings of the XIX International Congress of Psychology, London*.

———. 1979. "Smallest Space Analysis by the Absolute Value Principle." In *Geometric Representations of Relational Data*, edited by J. C. Lingoes, E. E. Roskam, and I. Borg, 707–12. Mathesis Press.

Hartmann, W. 1979. *Geometrische Modelle zur Analyse empirischer Data*. Akademie Verlag.

Hayashi, C. 1952. "On the Prediction of Phenomena from Qualitative Data and the Quantification of Qualitative Data from the Mathematico-Statistical Point of View." *Annals of the Institute of Statistical Mathematics* 3: 69–98.

Johnson, R. M. 1973. "Pairwise Nonmetric Multidimensional Scaling." *Psychometrika* 38 (12–18).

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.

Okamoto, M., and T. Isogai. 1878. "Optimality of Multidimensional Representation in Hayashi's Fourth Method of Quantification." *Journal of the Japanese Statistical Society* 8 (2): 63–69.

Roskam, E. E. 1970. "The Method of Triads for Nonmetric Multidimensional Scaling." *Nederlands Tijdschrift Voor de Psychologie* 25 (404-417).

———. 1979. "The Method of Triads for Nonmetric Multidimensional Scaling." In *Geometric Representations of Relational Data*, edited by J. C. Lingoes, E. E. Roskam, and I. Borg, 497–509. Mathesis Press.

Takane, Y. 1978a. "A Maximum Likelihood Method for Nonmetric Multidimensional Scaling: I. The Case in Which All Empirical Pairwise Orderings Are Independent - Evaluations." *Japanese Psychological Research* 20 (3): 105–14.

———. 1978b. "A Maximum Likelihood Method for Nonmetric Multidimensional Scaling: I. The Case in Which All Empirical Pairwise Orderings Are Independent - Theory." *Japanese Psychological Research* 20 (1): 7–17.

Torgerson, W. S. 1958. *Theory and Methods of Scaling*. New York: Wiley.