Powering Dissimilarities in Metric MDS

Jan de Leeuw - University of California Los Angeles

Started June 16 2023, Version of July 21, 2023

Abstract

For each integer p there is an r > 0 such that for all $s \ge r$ the Hadamard power Δ^s has Gower rank less than or equal to p.

Contents

1	Intr	oduction	2
2	Data	a	2
3	Resi	alts	3
	3.1	r = .25	4
	3.2	r = .50	5
	3.3	r=1	5
	3.4	r = 2	6
	3.5	r = 5	7
	3.6	r = 8	8
	3.7	r = 10	9
	3.8	r = 15	10
4	Disc	ussion	11
	4.1	Global pMDS	11
	4.2	Nonmetric MDS	11
	4.3	traceGower.R	12
	4.4	smacof.R	12
	4.5	smacofBasics.R	14
	4.6	janUtil.R	15
Re	eferen	ces	15

Note: This is a working paper which will be expanded/updated frequently. All suggestions for improvement are welcome.

1 Introduction

In Full Dimensional Multidimensional Scaling (FMDS) we minimize (metric, Euclidean) stress over all configurations of n points in n-1 dimensions. It was shown in De Leeuw (1993) that in FMDS there is a single local minimum, which is consequently global. There is a more detailed discussion of FMDS in De Leeuw, Groenen, and Mair (2016) and De Leeuw (2019a). In smacof notation (De Leeuw and Mair (2009), Mair, Groenen, and De Leeuw (2022)) a necessary and sufficient solution for X to be the global minimum of the FMDS problem is $X = V^+B(X)X$ and $V^+B(X) \leq I$ (all eigenvalues less than or equal to one). At the global minimum of FMDS the rank of X is called *the Gower Rank* of the dissimilarities (De Leeuw (2016)), which is also equal the number of eigenvalues of $V^+B(X)$ equal to one.

In p-dimensional multidimensional scaling (pMDS) we minimize stress over all $n \times p$ configurations, where generally p is much smaller than n - 1. The condition $X = V^+B(X)X$ remains necessary for a local minimum, but it is no longer sufficient, and in pMDS there can be (and usually are) many local minima. Note that the equation $X = V^+B(X)X$ means that $V^+B(X)$ has p eigenvalues equal to one. A pMDS solution is a global minimum if the largest eigenvalue of $V^+B(X)$ is equal to one. Actually finding the global minimum is a difficult computational problem, for which only approximate and somewhat tentative algorithms are available.

From the FMDS theory we know that if in pMDS solution $V^+B(X) \leq I$ then the pMDS solution has its global minimum at X, because in that case the FMDS and the pMDS solution are identical (De Leeuw (2014)). The resulting sufficient condition for a global minimum is, however, very seldom met in actual pMDS.

In De Leeuw (2019b) we use FMDS to approximate the global minimum of pMDS, using a quadratic penalty function that ultimate ensures that the final n - p dimensions of the FMDS solution are zero. So far, this has been rather successfull, although there are, of course, no guarantees.

In this paper we lay the foundation for a second way to globally minimize stress in pMDS using FMDS.

2 Data

The example in this paper uses the *wish* data from the smacof package, also analyzed in Kruskal and Wish (1978). The data are averages over 18 subjects of similarity judgments between 12 countries. We converted to dissimilarities by subtracting the similarity averages from seven. All weights in the stress loss function were set equal to one.

```
data(wish, package = "smacof")
countries <- as.matrix(wish)
w <- matrix(1, 12, 12) - diag(12)
countries <- 7 * w - countries
mPrint(countries, digits = 2)</pre>
```

##	BRAZIL	CONGO	CUBA	EGYPT	FRANCE	INDIA	ISRAEL
## BRAZIL	+0.00	+2.17	+1.72	+3.56	+2.28	+2.50	+3.17

##	CONGO	+2.17	+0.00	+2.44	+2.00	+3.00	+2.17	+3.67
##	CUBA	+1.72	+2.44	+0.00	+1.83	+2.89	+3.00	+3.39
##	EGYPT	+3.56	+2.00	+1.83	+0.00	+2.22	+1.17	+2.33
##	FRANCE	+2.28	+3.00	+2.89	+2.22	+0.00	+3.56	+3.00
##	INDIA	+2.50	+2.17	+3.00	+1.17	+3.56	+0.00	+2.89
##	ISRAEL	+3.17	+3.67	+3.39	+2.33	+3.00	+2.89	+0.00
##	JAPAN	+3.50	+3.61	+4.06	+3.17	+2.78	+2.50	+2.17
##	CHINA	+4.61	+3.00	+1.50	+2.61	+3.33	+2.89	+4.00
##	RUSSIA	+3.94	+3.61	+1.56	+2.61	+1.94	+2.50	+2.83
##	USA	+1.61	+4.61	+3.83	+3.67	+1.06	+2.72	+1.06
##	YUGOSLAV	+3.83	+3.50	+1.89	+2.72	+2.28	+3.00	+2.56
##		JAPAN	CHINA	RUSSIA	USA	YUGOSLAV		
##	BRAZIL	+3.50	+4.61	+3.94	+1.61	+3.83		
## ##	BRAZIL CONGO	+3.50 +3.61	+4.61 +3.00	+3.94 +3.61	+1.61 +4.61	+3.83 +3.50		
## ## ##	BRAZIL CONGO CUBA	+3.50 +3.61 +4.06	+4.61 +3.00 +1.50	+3.94 +3.61 +1.56	+1.61 +4.61 +3.83	+3.83 +3.50 +1.89		
## ## ## ##	BRAZIL CONGO CUBA EGYPT	+3.50 +3.61 +4.06 +3.17	+4.61 +3.00 +1.50 +2.61	+3.94 +3.61 +1.56 +2.61	+1.61 +4.61 +3.83 +3.67	+3.83 +3.50 +1.89 +2.72		
## ## ## ## ##	BRAZIL CONGO CUBA EGYPT FRANCE	+3.50 +3.61 +4.06 +3.17 +2.78	+4.61 +3.00 +1.50 +2.61 +3.33	+3.94 +3.61 +1.56 +2.61 +1.94	+1.61 +4.61 +3.83 +3.67 +1.06	+3.83 +3.50 +1.89 +2.72 +2.28		
## ## ## ## ## ##	BRAZIL CONGO CUBA EGYPT FRANCE INDIA	+3.50 +3.61 +4.06 +3.17 +2.78 +2.50	+4.61 +3.00 +1.50 +2.61 +3.33 +2.89	+3.94 +3.61 +1.56 +2.61 +1.94 +2.50	+1.61 +4.61 +3.83 +3.67 +1.06 +2.72	+3.83 +3.50 +1.89 +2.72 +2.28 +3.00		
## ## ## ## ## ##	BRAZIL CONGO CUBA EGYPT FRANCE INDIA ISRAEL	+3.50 +3.61 +4.06 +3.17 +2.78 +2.50 +2.17	+4.61 +3.00 +1.50 +2.61 +3.33 +2.89 +4.00	+3.94 +3.61 +1.56 +2.61 +1.94 +2.50 +2.83	+1.61 +4.61 +3.83 +3.67 +1.06 +2.72 +1.06	+3.83 +3.50 +1.89 +2.72 +2.28 +3.00 +2.56		
## ## ## ## ## ## ## ##	BRAZIL CONGO CUBA EGYPT FRANCE INDIA ISRAEL JAPAN	+3.50 +3.61 +4.06 +3.17 +2.78 +2.50 +2.17 +0.00	+4.61 +3.00 +1.50 +2.61 +3.33 +2.89 +4.00 +2.83	+3.94 +3.61 +1.56 +2.61 +1.94 +2.50 +2.83 +2.39	+1.61 +4.61 +3.83 +3.67 +1.06 +2.72 +1.06 +0.94	+3.83 +3.50 +1.89 +2.72 +2.28 +3.00 +2.56 +2.72		
 ## ## ## ## ## ## ## ##	BRAZIL CONGO CUBA EGYPT FRANCE INDIA ISRAEL JAPAN CHINA	+3.50 +3.61 +4.06 +3.17 +2.78 +2.50 +2.17 +0.00 +2.83	+4.61 +3.00 +1.50 +2.61 +3.33 +2.89 +4.00 +2.83 +0.00	+3.94 +3.61 +1.56 +2.61 +1.94 +2.50 +2.83 +2.39 +1.28	+1.61 +4.61 +3.83 +3.67 +1.06 +2.72 +1.06 +0.94 +4.44	+3.83 +3.50 +1.89 +2.72 +2.28 +3.00 +2.56 +2.72 +1.94		
 ## ## ## ## ## ## ## ## ##	BRAZIL CONGO CUBA EGYPT FRANCE INDIA ISRAEL JAPAN CHINA RUSSIA	+3.50 +3.61 +4.06 +3.17 +2.78 +2.50 +2.17 +0.00 +2.83 +2.39	+4.61 +3.00 +1.50 +2.61 +3.33 +2.89 +4.00 +2.83 +0.00 +1.28	+3.94 +3.61 +1.56 +2.61 +1.94 +2.50 +2.83 +2.39 +1.28 +0.00	+1.61 +4.61 +3.83 +3.67 +1.06 +2.72 +1.06 +0.94 +4.44 +2.00	+3.83 +3.50 +1.89 +2.72 +2.28 +3.00 +2.56 +2.72 +1.94 +0.33		
 ## ## ## ## ## ## ## ## ##	BRAZIL CONGO CUBA EGYPT FRANCE INDIA ISRAEL JAPAN CHINA RUSSIA USA	+3.50 +3.61 +4.06 +3.17 +2.78 +2.50 +2.17 +0.00 +2.83 +2.39 +0.94	+4.61 +3.00 +1.50 +2.61 +3.33 +2.89 +4.00 +2.83 +0.00 +1.28 +4.44	+3.94 +3.61 +1.56 +2.61 +1.94 +2.50 +2.83 +2.39 +1.28 +0.00 +2.00	+1.61 +4.61 +3.83 +3.67 +1.06 +2.72 +1.06 +0.94 +4.44 +2.00 +0.00	+3.83 +3.50 +1.89 +2.72 +2.28 +3.00 +2.56 +2.72 +1.94 +0.33 +3.44		

3 Results

We are interested in what happens to both FMDS and pMDS solutions (with p = 2) when we analyze dissimilarities that are transformed by raising them to different powers. Empirically we have found that the Gower rank of powered dissimilarities decreases as a function of the power r, commonly starting near n - 1 for r = 1 and ending up at one for large r. I emphasize this is not a theorem (yet), but just a computational finding in multiple examples.

In a sense, the result is not too surprising. For large r the dissimilarity matrix will be dominated by the largest dissimilarity, and the points with the largest dissimilarity can be placed at the endpoints of a line, with the rest in between. What is surprising, at least to me, is that for relatively low powers (say r < 10) we have a low Gower rank and still plenty of detail in the solution, which means that pMDS with p equal to the Gower rank finds the global minimum. In the color data from Ekman (1954), for example, raising the dissimilarities to the third power brings down the Gower rank to two (De Leeuw (2014)).

We give the results for a sequence of powers r on the Wish example, for now with very little comment. The smacof solutions all start with the classical Torgerson-Gower solution (Torgerson (1958)) and iterate until the stress values from one iteration to the next change by less than 1e - 15. That is ridiculously precise in an MDS context, but we want to make sure we are very close to a

local minimum.

3.1 r = .25

For small values of the power r the transformed dissimilarities will be close to the distances of a regular simplex. This means Gower rank will be n - 1. Minimum stress for FMDS will be very small, and for 2MDS it will be very large. This is what we find for r = .25. The two unit eigenvalues of $V^+B(X)$ are actually the smallest non-zero eigenvalues, indicating a very bad pMDS problem, with presumably a very large number of local minima.

```
## RESULTS FOR 0.2500
  [1] "minimum raw stress for FMDS and 2MDS"
##
## 0.0000000000000 2.084589217452540
  [1] "minimum normalized stress for FMDS and 2MDS"
##
## 0.0000000000000 0.020986853099850
## [1] "singular values X for FMDS"
    [1] +0.911823 +0.848901 +0.733064 +0.636770 +0.621680 +0.584975 +0.577622
##
##
    [8] +0.485632 +0.406526 +0.349673 +0.240971
##
  [1] "singular values X for 2MDS"
  [1] +1.394606 +1.358795
##
##
  [1] "eigenvalues V<sup>+</sup>B(X) for FMDS"
    [1] +1.000000 +1.000000 +1.000000 +1.000000 +1.000000 +1.000000
##
    [8] +1.000000 +1.000000 +1.000000 +0.000000
##
## [1] "eigenvalues V<sup>+</sup>B(X) for 2MDS"
##
    [1] +1.510326 +1.472451 +1.413206 +1.382267 +1.333851 +1.316327 +1.259150
    [8] +1.198866 +1.184617 +1.000000 +1.000000 +0.000000
##
```



The two plots show the 2MDS solution. The one on the left has a different scale for the two dimensions, the one on the right has both dimensions on the same scale. For small r the two plots are virtually the same.

3.2 r = .50

The same remarks as for r = .25 apply, only less extremely. The Gower rank is down to nine, but the unit eigenvalues are still the smallest ones.

```
## RESULTS FOR 0.5000
## [1] "minimum raw stress for FMDS and 2MDS"
## 0.016813189776687 1.176073277740564
## [1] "minimum normalized stress for FMDS and 2MDS"
## 0.000217318468206 0.015201306034834
## [1] "singular values X for FMDS"
##
  [1] +0.958007 +0.871687 +0.704340 +0.566183 +0.522508 +0.463092 +0.430679
    [8] +0.231413 +0.005028 +0.000000 +0.000000
##
## [1] "singular values X for 2MDS"
## [1] +1.257791 +1.202314
## [1] "eigenvalues V<sup>+</sup>B(X) for FMDS"
   [1] +1.000000 +1.000000 +1.000000 +1.000000 +1.000000 +1.000000
##
##
    [8] +1.000000 +0.999846 +0.983008 +0.965245 +0.000000
## [1] "eigenvalues V<sup>+</sup>B(X) for 2MDS"
##
    [1] +1.457879 +1.384164 +1.352360 +1.306342 +1.274894 +1.237496 +1.180982
    [8] +1.139662 +1.126206 +1.000000 +1.000000 +0.000000
##
```



3.3 r = 1

For the original untransformed data the Gower rank is six (or seven), and the unit eigenvalues of $V^+B(X)$ have moved up a place.

```
## RESULTS FOR 1.0000
## [1] "minimum raw stress for FMDS and 2MDS"
## 0.201887542090101 0.603629942647604
## [1] "minimum normalized stress for FMDS and 2MDS"
## 0.003992489316306 0.011937269987403
```

```
## [1] "singular values X for FMDS"
   [1] +0.912077 +0.805746 +0.579953 +0.389891 +0.288252 +0.144025 +0.000002
##
##
   [8] +0.000000 +0.000000 +0.000000 +0.000000
## [1] "singular values X for 2MDS"
## [1] +1.044677 +0.956555
## [1] "eigenvalues V^+B(X) for FMDS"
    [1] +1.000000 +1.000000 +1.000000 +1.000000 +1.000000 +0.995184
##
    [8] +0.959376 +0.918836 +0.898833 +0.849330 +0.000000
##
## [1] "eigenvalues V<sup>+</sup>B(X) for 2MDS"
   [1] +1.374950 +1.294842 +1.234365 +1.186110 +1.107626 +1.076704 +1.037704
##
##
    [8] +1.004519 +1.000000 +1.000000 +0.954487 +0.000000
```



3.4 r = 2

For r = 2 the Gower rank is down to four and the unit eigenvalues are up to fourth and fifth place.

```
## RESULTS FOR 2.0000
## [1] "minimum raw stress for FMDS and 2MDS"
## 0.600836884949776 0.660808460717618
## [1] "minimum normalized stress for FMDS and 2MDS"
## 0.022642989456655 0.024903063349961
## [1] "singular values X for FMDS"
##
   [1] +0.730826 +0.598564 +0.320623 +0.101539 +0.000000 +0.000000 +0.000000
   [8] +0.000000 +0.000000 +0.000000 +0.000000
##
## [1] "singular values X for 2MDS"
## [1] +0.766584 +0.638631
## [1] "eigenvalues V^+B(X) for FMDS"
## [1] +1.000000 +1.000000 +1.000000 +0.945852 +0.909598 +0.867725
   [8] +0.837378 +0.767077 +0.717658 +0.664230 +0.000000
##
## [1] "eigenvalues V^+B(X) for 2MDS"
## [1] +1.269220 +1.103698 +1.039430 +1.000000 +1.000000 +0.967035 +0.907976
##
    [8] +0.860553 +0.805728 +0.775803 +0.689188 +0.000000
```



3.5 r = 5

For r = 5 the Gower rank is down to three and the unit eigenvalues are in second and third place. The first two dimensions of FMDS have stress very close to stress of the pMDS solution. We now also see from the two plots that the second dimension of the configuration is collapsing.

```
## RESULTS FOR 5.0000
## [1] "minimum raw stress for FMDS and 2MDS"
## 0.889942897327712 0.890439775666370
## [1] "minimum normalized stress for FMDS and 2MDS"
## 0.094053041997496 0.094105554264737
## [1] "singular values X for FMDS"
##
    [1] +0.437804 +0.228253 +0.046478 +0.000000 +0.000000 +0.000000 +0.000000
##
    [8] +0.000000 +0.000000 +0.000000 +0.000000
## [1] "singular values X for 2MDS"
##
  [1] +0.438715 +0.231037
  [1] "eigenvalues V<sup>+</sup>B(X) for FMDS"
##
    [1] +1.000000 +1.000000 +0.815716 +0.752369 +0.639559 +0.587647
##
    [8] +0.531858 +0.516145 +0.458722 +0.387285 +0.000000
##
## [1] "eigenvalues V<sup>+</sup>B(X) for 2MDS"
##
    [1] +1.072992 +1.000000 +1.000000 +0.824166 +0.759072 +0.645328 +0.588648
    [8] +0.540184 +0.518484 +0.458735 +0.388575 +0.000000
##
```



3.6 r = 8

For r = 8 the Gower rank is down to two and the unit eigenvalues are in first and second place. The 2MDS and FMDS solutions are identical, both equal to the global minimum of pMDS for dissimilarities to the power eight. The second dimension is now almost non-existent in the plot on the right.

```
## RESULTS FOR 8.0000
## [1] "minimum raw stress for FMDS and 2MDS"
## 0.861528292644553 0.861528292644499
## [1] "minimum normalized stress for FMDS and 2MDS"
## 0.136700436161952 0.136700436161944
## [1] "singular values X for FMDS"
##
    [1] +0.344493 +0.018232 +0.000001 +0.000000 +0.000000 +0.000000 +0.000000
    [8] +0.000000 +0.000000 +0.000000 +0.000000
##
##
  [1] "singular values X for 2MDS"
## [1] +0.344493 +0.018232
## [1] "eigenvalues V^+B(X) for FMDS"
    [1] +1.000000 +1.000000 +0.988788 +0.809165 +0.576103 +0.543085 +0.472001
##
    [8] +0.388677 +0.344025 +0.307488 +0.239859 +0.000000
##
## [1] "eigenvalues V^+B(X) for 2MDS"
    [1] +1.000000 +1.000000 +0.988788 +0.809165 +0.576102 +0.543084 +0.472001
##
    [8] +0.388676 +0.344025 +0.307488 +0.239859 +0.000000
##
```



3.7 r = 10

RESULTS FOR 10.0000 ## [1] "minimum raw stress for FMDS and 2MDS" ## 0.849049190259466 0.849049190259466 ## [1] "minimum normalized stress for FMDS and 2MDS" **##** 0.153489609905732 0.153489609905732 ## [1] "singular values X for FMDS" [1] +0.298222 +0.006342 +0.000000 +0.000000 +0.000000 +0.000000 +0.000000 ## [8] +0.000000 +0.000000 +0.000000 +0.000000 ## ## [1] "singular values X for 2MDS" ## [1] +0.298222 +0.006342 ## [1] "eigenvalues V^+B(X) for FMDS" [1] +1.000000 +1.000000 +0.911000 +0.785987 +0.475049 +0.418898 +0.357726 ## [8] +0.298582 +0.250702 +0.201013 +0.158641 +0.000000 ## ## [1] "eigenvalues V^+B(X) for 2MDS" [1] +1.000001 +1.000000 +0.911000 +0.785987 +0.475049 +0.418898 +0.357726 ## [8] +0.298583 +0.250702 +0.201013 +0.158641 +0.000000



3.8 r = 15

RESULTS FOR 15.0000 ## [1] "minimum raw stress for FMDS and 2MDS" ## 0.838657370199356 0.838657370199360 ## [1] "minimum normalized stress for FMDS and 2MDS" ## 0.176036341301155 0.176036341301156 ## [1] "singular values X for FMDS" ## [1] +0.242340 +0.000000 +0.000000 +0.000000 +0.000000 +0.000000 +0.000000 ## [8] +0.000000 +0.000000 +0.000000 +0.000000 ## [1] "singular values X for 2MDS" ## [1] +0.242340 +0.000000 ## [1] "eigenvalues V^+B(X) for FMDS" [1] +1.000000 +0.927400 +0.849990 +0.794779 +0.273187 +0.241039 +0.169889 ## ## [8] +0.130252 +0.101554 +0.072569 +0.042251 +0.000000 ## [1] "eigenvalues V^+B(X) for 2MDS" [1] +1.000000 +0.927400 +0.849990 +0.794779 +0.273187 +0.241039 +0.169889 ## [8] +0.130252 +0.101554 +0.072569 +0.042251 -0.000000



4 Discussion

4.1 Global pMDS

Our results suggest a possible global minimization method for pMDS, based on FMDS. First find a power for which the Gower rank is p, generally by increasing r starting at one. Then use the corresponding pMDS solution as starting point of a sequence of pMDS solutions in which r is decreasing to one. In each of these solutions we use the solution of the previous pMDS problem as the starting point of the iterations. This defines a trajectory which hopefully gets us to a good local minimum. We have not tried this strategy yet, and of course it is more likely to work if Gower rank equal to p happens at a fairly low power of r.

It should perhaps be mentioned here that data which are averages of similarity or dissimilarity judgments over individuals in a heterogenous group tend to regress to the mean, and are consequently biased to produce solutions which look like projections of a regular simplex on a plane. This is true for the Wish data and also for the Ekman data we mentioned before (and for many of the data sets analyzed in De Leeuw (2019b)). A transformation which stretches the scale at the endpoints is often called for.

4.2 Nonmetric MDS

The results in this paper are relevant for nonmetric pMDS, because the power transform is an increasing function for any positive power. Thus each of the solutions in the results section of this paper are possible nonmetric pMDS solutions. In this context it is important to observe that normalized stress generally increases with increasing power r (again, not a theorem).

The fact that we use power transformations to study Gower rank is, to some extent, arbitrary. Any parametric family of monotonic transformations would be of interest. Our results show that in nonmetric pMDS there are many transformations of the dissimilarities with Gower rank equal to p, and for such transformed data a metric pMDS finds the global minimum. It remains to find out what an actual nonmetric pMDS actually does to the Gower rank. But we must keep in mind that there

is no such thing as nonmetric FMDS, because there is always a monotone transformation of the dissimilarities that can be perfectly fitted by Euclidean distances in n - 2 dimensions (De Leeuw (1970)). Thus nonmetric FMDS always has minimum stress equal to zero. # Appendix: Code

4.3 traceGower.R

```
traceGower <- function(data, power) {</pre>
  pful <- nrow(data) - 1
  data <- data ^ power
  data <- data / max(data)</pre>
  deno <- sum(data ^ 2)
  h <- smacof(w, data, p = pful, eps = 1e-15)</pre>
  g \leftarrow smacof(w, data, p = 2, eps = 1e-15)
  cat("RESULTS FOR", formatC(power, digits = 4, format = "f"), "\n")
  print("minimum raw stress for FMDS and 2MDS")
  cat(formatC(c(h$s , g$s), digits = 15, format = "f"), "\n")
  print("minimum normalized stress for FMDS and 2MDS")
  cat(formatC(c(h$s / deno, g$s / deno), digits = 15, format = "f"), "\n")
  print("singular values X for FMDS")
  mPrint(svd(h$x)$d, d = 6)
  print("singular values X for 2MDS")
  mPrint(svd(g$x)$d, d = 6)
  print("eigenvalues V^+B(X) for FMDS")
  mPrint((eigen(h$b)$values)/12, d = 6)
  print("eigenvalues V^+B(X) for 2MDS")
  mPrint((eigen(g$b)$values)/12, d = 6)
  par(pty = "s")
  lmax <- max(g$x)</pre>
  lmin <- min(g$x)</pre>
  plot(g$x, type = "n", xlab = "dimension 1", ylab = "dimension 2")
  text(g$x, row.names(countries))
  plot(g$x, type = "n", xlim = c(lmin,lmax), ylim = c(lmin,lmax), xlab = "dimension 1",
  text(g$x, row.names(countries))
}
```

4.4 smacof.R

```
p = 2,
       xold = torgerson(delta, p),
       itmax = 10000,
       eps = 1e-10,
       verbose = FALSE) {
itel <- 1
n <- nrow(xold)</pre>
vmat <- smacofVmat(w)</pre>
vinv <- solve(vmat + (1 / n)) - (1 / n)</pre>
dold <- as.matrix(dist(xold))</pre>
mold <- sum(w * delta * dold) / sum(w * dold * dold)</pre>
xold <- xold * mold</pre>
dold <- dold * mold
sold <- smacofLoss(dold, w, delta)</pre>
bold <- smacofBmat(dold, w, delta)</pre>
repeat {
  xnew <- smacofGuttman(xold, bold, vinv)</pre>
  dnew <- as.matrix(dist(xnew))</pre>
  bnew <- smacofBmat(dnew, w, delta)</pre>
  snew <- smacofLoss(dnew, w, delta)</pre>
  if (verbose) {
    cat(
      "itel ",
      formatC(itel, width = 4, format = "d"),
      "sold ",
      formatC(
        sold,
        width = 10,
        digits = 6,
        format = "f"
      ),
      "snew ",
      formatC(
        snew,
        width = 10,
        digits = 6,
        format = "f"
      ),
     "\n"
    )
  }
  if (((sold - snew) < eps) || (itel == itmax)) {</pre>
    break
  }
```

```
itel <- itel + 1
xold <- xnew
bold <- bnew
sold <- snew
}
return(list(
    x = xnew,
    d = dnew,
    b = bnew,
    s = snew,
    itel = itel
))
}</pre>
```

4.5 smacofBasics.R

```
smacofLoss <- function(d, w, delta) {</pre>
  return(sum(w * (delta - d) ^ 2) / 4)
}
smacofBmat <- function(d, w, delta) {</pre>
  dd <- ifelse(d == 0, 0, 1 / d)
  b <- -dd * w * delta
  diag(b) <- -rowSums(b)</pre>
  return(b)
}
smacofVmat <- function(w) {</pre>
  v <- −w
  diag(v) <- -rowSums(v)
  return(v)
}
smacofGuttman <- function(x, b, vinv) {</pre>
  return(vinv %*% b %*% x)
}
doubleCenter <- function(x) {</pre>
  rs <- apply(x, 1, mean)</pre>
 ss <- mean(x)
  return(x - outer(rs, rs, "+") + ss)
}
```

```
columnCenter <- function(x) {
  return(apply(x, 2, function(z) z - mean(z)))
}
torgerson <- function(delta, p) {
  e <- eigen(-.5 * doubleCenter(delta ^ 2))
  l <- sqrt(pmax(0, e$values[1:p]))
  return(e$vectors[, 1:p] %*% diag(l))
}</pre>
```

4.6 janUtil.R

```
mPrint <- function(x,</pre>
                     digits = 6,
                     width = 8,
                     format = "f",
                     flag = "+") {
  print(noquote(
    formatC(
      x,
      digits = digits,
      width = width,
      format = format,
      flag = flag
    )
  ))
}
butLast <- function(x, m = 1) {</pre>
  return(rev(rev(x)[-(1:m)]))
}
butFirst <- function(x, m = 1) {</pre>
  return(x[-(1:m)])
}
```

References

De Leeuw, J. 1970. "The Euclidean Distance Model." Research Note 002-70. Department of Data Theory FSW/RUL.

—. 1993. "Fitting Distances by Least Squares." Preprint Series 130. Los Angeles, CA: UCLA Department of Statistics. https://jansweb.netlify.app/publication/deleeuw-r-93-c/deleeuw-r-93-c.pdf.

—. 2014. "Bounding, and Sometimes Finding, the Global Minimum in Multidimensional Scaling." UCLA Department of Statistics. https://jansweb.netlify.app/publication/deleeuw-u-14-b/deleeuw-u-14-b.pdf.

——. 2016. "Gower Rank." 2016. https://jansweb.netlify.app/publication/deleeuw-e-16-k/deleeuw-e-16-k.pdf.

——. 2019a. "Generalized Full-dimensional Scaling." 2019. https://jansweb.netlify.app/publicat ion/deleeuw-e-19-f/deleeuw-e-19-f.pdf.

- ——. 2019b. "Global Minima by Penalized Full-dimensional Scaling." 2019. https://jansweb.ne tlify.app/publication/deleeuw-e-19-e/deleeuw-e-19-e.pdf.
- De Leeuw, J., P. Groenen, and P. Mair. 2016. "Full-Dimensional Scaling." 2016. https://jansweb. netlify.app/publication/deleeuw-groenen-mair-e-16-e/deleeuw-groenen-mair-e-16-e.pdf.
- De Leeuw, J., and P. Mair. 2009. "Multidimensional Scaling Using Majorization: SMACOF in R." *Journal of Statistical Software* 31 (3): 1–30. https://www.jstatsoft.org/article/view/v031i03.
- Ekman, G. 1954. "Dimensions of Color Vision." Journal of Psychology 38: 467-74.

Kruskal, J. B., and M. Wish. 1978. Multidimensional Scaling. Sage.

Mair, P., P. J. F. Groenen, and J. De Leeuw. 2022. "More on Multidimensional Scaling in R: smacof Version 2." *Journal of Statistical Software* 102 (10): 1–47. https://www.jstatsoft.org/article/vi ew/v102i10.

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