

# Least Squares Euclidean Multidimensional Scaling

Jan de Leeuw

Started October 02, 2020. Last update January 30, 2024



# Contents

<b>Note</b>	<b>19</b>
<b>Preface</b>	<b>23</b>
<b>Notation and Reserved Symbols</b>	<b>27</b>
0.1 Spaces . . . . .	27
0.2 Matrices . . . . .	28
0.3 Functions . . . . .	28
0.4 MDS . . . . .	28
<b>1 Introduction</b>	<b>31</b>
1.1 Brief History . . . . .	32
1.1.1 Milestones . . . . .	32
1.2 Basic MDS . . . . .	32
1.2.1 Kruskal's stress . . . . .	32
1.2.1.1 Square root . . . . .	32
1.2.1.2 Weights . . . . .	32
1.2.1.3 Normalization . . . . .	32
1.3 Local and Global . . . . .	32
1.4 Generalizations . . . . .	32
1.4.1 Non-metric MDS . . . . .	32

1.4.2	fstress . . . . .	32
1.4.3	Constraints . . . . .	32
1.4.4	Individual Differences . . . . .	32
1.4.5	Distance Asymmetry . . . . .	32
1.5	Models and Techniques . . . . .	32
<b>2</b>	<b>Properties of Stress</b>	<b>33</b>
2.1	Notation . . . . .	35
2.1.1	Expanding . . . . .	35
2.1.2	Matrix Expressions . . . . .	35
2.1.3	Coefficient Space . . . . .	35
2.1.4	Our Friends CS and AM/GM . . . . .	35
2.2	Global Properties . . . . .	35
2.2.1	Boundedness . . . . .	35
2.2.2	Invariance . . . . .	35
2.2.3	Continuity . . . . .	35
2.2.4	Coercivity . . . . .	35
2.3	Differentiability . . . . .	35
2.3.0.1	Distances . . . . .	35
2.3.0.2	Rho and Stress . . . . .	35
2.3.0.3	expandStress . . . . .	35
2.3.1	Partial Derivatives . . . . .	35
2.3.2	Special Expansions . . . . .	35
2.3.2.1	Infinitesimal Rotations . . . . .	35
2.3.2.2	Singularities . . . . .	35
2.3.2.3	Singularities . . . . .	35
2.4	Convexity . . . . .	35

<i>CONTENTS</i>	5
-----------------	---

2.4.1	Distances . . . . .	35
2.4.2	Subdifferentials . . . . .	35
2.4.3	DC Functions . . . . .	35
2.4.4	Negative Dissimilarities . . . . .	35
2.5	Stationary Points . . . . .	35
2.5.1	Local Maxima . . . . .	35
2.5.2	Local Minima . . . . .	35
2.5.3	Saddle Points . . . . .	35
2.5.4	An Example . . . . .	35
2.5.4.1	Regular Tetrahedron . . . . .	35
2.5.4.2	Singularity . . . . .	35
2.5.4.3	Equilateral Triangle with Centroid . . . . .	35
2.5.4.4	Square . . . . .	35
2.5.4.5	Non-global Local Minima . . . . .	35
2.5.4.6	Directions of Descent . . . . .	35

### **3 Stress Spaces 37**

3.1	Configuration Space . . . . .	39
3.1.1	Zero Distance Subspaces . . . . .	39
3.2	Spherical Space . . . . .	39
3.3	Distance Space . . . . .	39
3.4	Squared Distance Space . . . . .	39
3.5	Gramian Space . . . . .	39
3.6	Pictures of Stress . . . . .	39
3.6.1	Coefficient Space . . . . .	39
3.6.2	Global Perspective . . . . .	39
3.6.3	Global Contour . . . . .	39

3.6.4	Stationary Points . . . . .	39
3.6.4.1	First Minimum . . . . .	39
3.6.5	Second Minimum . . . . .	39
3.6.6	Third Minimum . . . . .	39
3.6.7	First Saddle Point . . . . .	39
3.6.8	Second Saddle Point . . . . .	39
3.7	Another Look . . . . .	39
3.8	A Final Look . . . . .	39
3.9	Discussion . . . . .	39
3.10	Coordinates . . . . .	39
<b>4</b>	<b>Classical Multidimensional Scaling</b>	<b>41</b>
4.1	Algebra . . . . .	41
4.1.1	Torgerson Transform . . . . .	41
4.1.2	Schoenberg's Theorem . . . . .	41
4.2	Approximation . . . . .	41
4.2.1	Low Rank Approximation of the Torgersen Transform .	41
4.2.2	Minimization of Strain . . . . .	41
4.2.3	Approximation from Below . . . . .	41
<b>5</b>	<b>Minimization of Basic Stress</b>	<b>43</b>
5.1	Gradient Methods . . . . .	45
5.1.1	Step Size . . . . .	45
5.1.1.1	Kruskal Step Size . . . . .	45
5.1.1.2	Guttman Step Size . . . . .	45
5.1.1.3	Cauchy Step Size . . . . .	45
5.1.1.4	Majorization Step Size . . . . .	45
5.2	Initial Configurations . . . . .	45

5.3	On MM Algorithms . . . . .	45
5.4	Smacof Algorithm . . . . .	45
5.4.1	Guttman Transform . . . . .	45
5.4.2	Subdifferentials . . . . .	45
5.4.3	Derivative . . . . .	45
5.4.4	Global Convergence . . . . .	45
5.4.4.1	From the CS Inequality . . . . .	45
5.4.4.2	From Majorization . . . . .	45
5.4.4.3	From Ratio of Norms . . . . .	45
5.4.5	Component Rotated Smacof . . . . .	45
5.4.6	Local Convergence . . . . .	45
5.4.6.1	Cross Product Iterations . . . . .	45
5.4.6.2	Rotation to PC . . . . .	45
5.4.7	Data Asymmetry . . . . .	45
5.4.8	Replications . . . . .	45
5.4.9	Negative Dissimilarities . . . . .	45
5.4.10	Normalization . . . . .	45
5.4.11	Unweighting . . . . .	45
5.4.11.1	Symmetric non-negative matrix factorization . . . . .	45
5.5	Stress Envelopes . . . . .	45
5.5.1	CS Majorization . . . . .	45
5.5.2	AM/GM Minorization . . . . .	45
5.5.3	Dualities . . . . .	45
5.6	Smacof in Coefficient Space . . . . .	45
5.7	Newton in MDS . . . . .	45
5.7.1	Regions of Attraction . . . . .	45
5.7.1.1	Smacof . . . . .	45

5.7.1.2	Newton . . . . .	45
5.8	Distance Smoothing . . . . .	45
<b>6</b>	<b>Acceleration of Convergence</b>	<b>47</b>
6.1	Simple Acceleration . . . . .	47
6.2	One-Parameter Methods . . . . .	47
6.3	SQUAREM . . . . .	47
6.4	Vector Extrapolation Methods . . . . .	47
<b>7</b>	<b>Nonmetric MDS</b>	<b>49</b>
7.1	Generalities . . . . .	50
7.1.1	Kruskal's Stress . . . . .	50
7.2	Single and Double Phase . . . . .	50
7.2.1	Double Phase . . . . .	50
7.2.2	Single Phase . . . . .	50
7.3	Affine NMDS . . . . .	50
7.4	Conic NMDS . . . . .	50
7.4.1	Normalization . . . . .	50
7.4.2	Normalized Cone Regression . . . . .	50
7.4.3	Hard Squeeze and Soft Squeeze . . . . .	50
7.4.4	Inner Iterations . . . . .	50
7.4.5	Stress-1 and Stress-2 . . . . .	50
<b>8</b>	<b>Interval MDS</b>	<b>51</b>
8.1	The Additive Constant . . . . .	52
8.1.1	Early . . . . .	52
8.1.2	Cooper . . . . .	52
8.2	Algebra . . . . .	52



<i>CONTENTS</i>	9
8.2.1 Existence . . . . .	52
8.2.2 Solution . . . . .	52
8.2.3 Examples . . . . .	52
8.2.3.1 Small Example . . . . .	52
8.2.3.2 De Gruijter Example . . . . .	52
8.2.3.3 Ekman Example . . . . .	52
8.2.4 A Variation . . . . .	52
8.3 Interval smacof . . . . .	52
8.3.1 Example . . . . .	52
<b>9 Polynomial MDS</b>	<b>53</b>
9.1 Introduction . . . . .	54
9.2 Fitting Polynomials . . . . .	54
9.3 Positive and Convex, Monotone Polynomials . . . . .	54
9.3.1 Introduction . . . . .	54
9.3.2 A QP Algorithm . . . . .	54
9.3.2.1 Example 1: Simple Monotone Regression . . .	54
9.3.2.2 Example 2: Monotone Regression with Ties .	54
9.3.2.3 Example 3: Weighted Rounding . . . . .	54
9.3.2.4 Example 4: Monotone Polynomials . . . . .	54
9.3.3 Examples . . . . .	54
<b>10 Ordinal MDS</b>	<b>55</b>
10.1 Monotone Regression . . . . .	56
10.1.1 Simple Monotone Regression . . . . .	56
10.1.2 Weighted Monotone Regression . . . . .	56
10.1.3 Normalized Cone Regression . . . . .	56
10.1.4 Iterative MR . . . . .	56

10.2 Alternating Least Squares . . . . .	56
10.3 Kruskal's Approach . . . . .	56
10.3.1 Kruskal's Stress . . . . .	56
10.3.2 Stress1 and Stress2 . . . . .	56
10.4 Guttman's Approach . . . . .	56
10.4.0.1 Rank Images . . . . .	56
10.4.0.2 Single and Double Phase . . . . .	56
10.4.0.3 Hard and Soft Squeeze . . . . .	56
10.4.1 Smoothness of Ordinal Loss Functions . . . . .	56
10.5 Scaling with Distance Bounds . . . . .	56
10.6 Bounds on Stress . . . . .	56
<b>11 Splinal MDS</b>	<b>57</b>
11.1 Splines . . . . .	58
11.1.1 B-splines . . . . .	58
11.1.1.1 Boundaries . . . . .	58
11.1.1.2 Normalization . . . . .	58
11.1.1.3 Recursion . . . . .	58
11.1.1.4 Illustrations . . . . .	58
11.1.2 I-splines . . . . .	58
11.1.2.1 Increasing Coefficients . . . . .	58
11.1.2.2 Increasing Values . . . . .	58
11.1.3 Time Series Example . . . . .	58
11.1.3.1 B-splines . . . . .	58
11.1.3.2 I-splines . . . . .	58
11.1.3.3 B-Splines with monotone weights . . . . .	58
11.1.3.4 B-Splines with monotone values . . . . .	58
11.1.4 Local positivity, monotonicity, convexity . . . . .	58

<i>CONTENTS</i>	11
<b>12 Unidimensional Scaling</b>	<b>59</b>
12.1 An example . . . . .	60
12.1.1 Perspective . . . . .	60
12.1.2 Contour . . . . .	60
12.2 Order Formulation . . . . .	60
12.3 Permutation Formulation . . . . .	60
12.4 Sign Matrix Formulation . . . . .	60
12.5 Algorithms for UMDS . . . . .	60
12.5.1 SMACOF . . . . .	60
12.5.2 SMACOF (smoothed) . . . . .	60
12.5.3 Branch-and-Bound . . . . .	60
12.5.4 Dynamic Programming . . . . .	60
12.5.5 Simulated Annealing . . . . .	60
12.5.6 Penalty Methods . . . . .	60
<b>13 Full-dimensional Scaling</b>	<b>61</b>
13.1 Convexity . . . . .	61
13.2 Optimality . . . . .	61
13.3 Iteration . . . . .	61
13.4 Cross Product Space . . . . .	61
13.5 Full-dimensional Scaling . . . . .	61
13.6 Ekman example . . . . .	61
<b>14 Unfolding</b>	<b>63</b>
14.1 Algebra . . . . .	63
14.1.1 One-dimensional . . . . .	63
14.2 Classical Unfolding . . . . .	63
14.3 Nonmetric Unfolding . . . . .	63

14.3.1	Degenerate Solutions . . . . .	63
14.3.1.1	Which Stress . . . . .	63
14.3.1.2	l'Hôpital's Rule . . . . .	63
14.3.1.3	Penalizing . . . . .	63
14.3.1.4	Restricting Regression . . . . .	63
<b>15</b>	<b>Constrained Multidimensional Scaling</b>	<b>65</b>
15.1	Primal-Dual (note: the base partitioning has dual aspects) . .	67
15.2	Basic Partitioning . . . . .	67
15.3	Unweighting . . . . .	67
15.4	Constraints on the Distances . . . . .	67
15.4.1	Rectangles . . . . .	67
15.5	Linear Constraints . . . . .	67
15.5.1	Uniqueness . . . . .	67
15.5.2	Combinations . . . . .	67
15.5.3	Step Size . . . . .	67
15.5.4	Single Design Matrix . . . . .	67
15.5.5	Multiple Design Matrices . . . . .	67
15.6	Circular MDS . . . . .	67
15.6.1	Some History . . . . .	67
15.6.2	Primal Methods . . . . .	67
15.6.3	Dual Methods . . . . .	67
15.7	Elliptical MDS . . . . .	67
15.7.1	Primal . . . . .	67
15.7.2	Dual . . . . .	67
15.8	Distance Bounds . . . . .	67
15.9	Localized MDS . . . . .	67

<i>CONTENTS</i>	13
15.10 MDS as MVA . . . . .	67
15.11 Horseshoes . . . . .	67
<b>16 Individual Differences</b>	<b>69</b>
16.1 Stress Majorization . . . . .	70
16.2 Types of Constraints . . . . .	70
16.2.1 Unconstrained . . . . .	70
16.2.2 Replications . . . . .	70
16.2.3 INDSCAL/PARAFAC . . . . .	70
16.2.4 IDIOSCAL/TUCKALS . . . . .	70
16.2.5 PARAFAC2 . . . . .	70
16.2.6 Factor Models . . . . .	70
16.3 Nonmetric Individual Differences . . . . .	70
16.3.1 Conditionality . . . . .	70
<b>17 Asymmetry in MDS</b>	<b>71</b>
17.1 Conditional Rankings . . . . .	71
17.2 Confusion Matrices . . . . .	71
17.3 The Slide Vector . . . . .	71
17.4 DEDICOM . . . . .	71
17.5 Constantine-Gower . . . . .	71
<b>18 Nominal MDS</b>	<b>73</b>
18.1 Binary Dissimilarities . . . . .	73
18.2 Indicator matrix . . . . .	73
18.3 Unfolding Indicator Matrices . . . . .	73
18.4 Linear Separation . . . . .	73
18.5 Circular Separation . . . . .	73

18.6 Convex Hull Scaling . . . . .	73
18.7 Voronoi Scaling . . . . .	73
18.8 Multidimensional Scalogram Analysis . . . . .	73
<b>19 Nonmonotonic MDS</b>	<b>75</b>
19.1 Filler . . . . .	75
<b>20 Compound Objects</b>	<b>77</b>
20.1 Filler . . . . .	77
<b>21 Sstress and strain</b>	<b>79</b>
21.1 sstress . . . . .	80
21.1.1 sstress and stress . . . . .	80
21.1.2 Decomposition . . . . .	80
21.1.3 Full-dimensional sstress . . . . .	80
21.1.4 Minimizing sstress . . . . .	80
21.1.4.1 ALSCAL . . . . .	80
21.1.4.2 Majorization . . . . .	80
21.1.4.3 SOS . . . . .	80
21.1.4.4 Duality . . . . .	80
21.1.4.5 ALS Unfolding . . . . .	80
21.1.5 Bounds for sstress . . . . .	80
21.2 strain . . . . .	80
21.2.1 Unweighted . . . . .	80
21.2.2 Bailey-Gower . . . . .	80
21.2.3 Using Additivity . . . . .	80

<i>CONTENTS</i>	15
<b>22 fstress and rstress</b>	<b>81</b>
22.1 fstress . . . . .	81
22.1.1 Use of Weights . . . . .	81
22.1.2 Convexity . . . . .	81
22.2 rStress . . . . .	81
22.2.1 Using Weights . . . . .	81
22.2.2 Minimizing rstress . . . . .	81
22.3 mstress . . . . .	81
22.4 astress . . . . .	81
22.5 pstress . . . . .	81
<b>23 Alternative Least Squares Loss</b>	<b>83</b>
23.1 Sammon's MDS . . . . .	83
23.2 Kamade-Kawai Spring . . . . .	83
23.3 McGee's Work . . . . .	83
23.4 Shepard's Nonmetric MDS . . . . .	83
23.5 Guttman's Nonmetric MDS . . . . .	83
23.6 Positive Orthant Nonmetric MDS . . . . .	83
23.7 Role Reversal . . . . .	84
<b>24 Inverse Multidimensional Scaling</b>	<b>85</b>
24.1 Basic IMDS . . . . .	87
24.2 Non-negative Dissimilarities . . . . .	87
24.3 Zero Weights and/or Distances . . . . .	87
24.4 Examples . . . . .	87
24.4.1 First Example . . . . .	87
24.4.2 Second Example . . . . .	87
24.4.3 Third Example . . . . .	87

24.4.4 Fourth Example . . . . .	87
24.5 MDS Sensitivity . . . . .	87
24.6 Second Order Inverse MDS . . . . .	87
24.7 Inverse FDS . . . . .	87
24.8 Multiple Solutions . . . . .	87
24.9 Minimizing iStress . . . . .	87
24.10 Order three . . . . .	87
24.11 Order Four . . . . .	87
24.12 Sensitivity . . . . .	87
<b>25 Stability of MDS Solutions</b>	<b>89</b>
25.1 Null Distribution . . . . .	89
25.2 Pseudo-confidence Ellipsoids . . . . .	89
25.3 A Pseudo-Bootstrap . . . . .	89
25.4 How Large is My Stress ? . . . . .	89
<b>26 In Search of Global Minima</b>	<b>91</b>
26.1 Random Starts . . . . .	93
26.1.1 Examples . . . . .	93
26.1.1.1 Ekman . . . . .	93
26.1.1.2 De Gruijter . . . . .	93
26.2 Tunneling, Filling, Annealing, etc. . . . .	93
26.3 Cutting Planes . . . . .	93
26.3.1 On the Circle . . . . .	93
26.3.2 Cauchy Step Size . . . . .	93
26.3.3 Balls . . . . .	93
26.3.3.1 Outer Approximation . . . . .	93
26.4 Distance Smoothing . . . . .	93



<i>CONTENTS</i>	17
-----------------	----

26.5 Penalizing Dimensions . . . . .	93
26.5.1 Local Minima . . . . .	93
26.5.2 Algorithm . . . . .	93
26.5.3 Examples . . . . .	93
26.5.3.1 Chi Squares . . . . .	93
26.5.3.2 Regular Simplex . . . . .	93
26.5.3.3 Intelligence . . . . .	93
26.5.3.4 Countries . . . . .	93
26.5.3.5 Dutch Political Parties . . . . .	93
26.5.3.6 Ekman . . . . .	93
26.5.3.7 Morse in Two . . . . .	93
26.5.3.8 Vegetables . . . . .	93
26.5.3.9 Plato . . . . .	93
26.5.3.10 Morse in One . . . . .	93

<b>27 Software</b>	<b>95</b>
--------------------	-----------

.1 R Code . . . . .	97
.1.1 utilities.R . . . . .	97
.1.2 common/indexing.r . . . . .	97
.1.3 common/io.r . . . . .	97
.1.4 common/linear.r . . . . .	97
.1.5 common/nextPC.r . . . . .	97
.1.6 common/smacof.r . . . . .	97
.1.7 properties.R . . . . .	97
.1.8 expandOneDim.R . . . . .	97
.1.9 pictures.R . . . . .	97
.1.10 classical.R . . . . .	97

.1.11	minimization.R . . . . .	97
.1.12	full.R . . . . .	97
.1.13	unfolding.R . . . . .	97
.1.14	constrained.R . . . . .	97
.1.15	nominal.R . . . . .	97
.1.16	sstress.R . . . . .	97
.1.17	inverse.R . . . . .	97
.1.18	global.R . . . . .	97
.1.19	mathadd.R . . . . .	97
.2	C code . . . . .	97
.2.1	deboor.c . . . . .	97
.2.2	cleanup.c . . . . .	97
.2.3	jacobi.c . . . . .	97
.2.4	jbkTies.c . . . . .	97
.2.5	matrix.c . . . . .	97
.2.6	jeffrey.c . . . . .	97
.2.7	smacofBlockSort.c . . . . .	97
.2.8	smacofConvert.c . . . . .	97
.2.9	nextPC.c . . . . .	97
.3	Small . . . . .	97
.4	De Gruijter . . . . .	97
.5	Ekman . . . . .	97
.6	Vegetables . . . . .	97

# Note

This book will be expanded/updated frequently. The directory [github.com/deleeuw/stress](https://github.com/deleeuw/stress) has a pdf version, a html version, the bib file, the complete Rmd file with the codechunks, and the R and C source code. All suggestions for improvement of text or code are welcome, and some would be really beneficial. For example, I only use base R graphics, nothing more fancy, because base graphics is all I know.

All text and code are in the public domain and can be copied, modified, and used by anybody in any way they see fit. Attribution will be appreciated, but is not required. For completeness we include a slightly modified version of the Unlicense as appendix ??.

I number and label *all* displayed equations. Equations are displayed, instead of inlined, if and only if one of the following is true.

- They are important.
- They are referred to elsewhere in the text.
- Not displaying them messes up the line spacing.

All code chunks in the text are named. Theorems, lemmas, chapters, sections, subsections and so on are also named and numbered, using book-down/Rmarkdown.

I have been somewhat hesitant to use lemmas, theorems, and corollaries in this book. But ultimately they enforce precision and provide an excellent organizational tool. If there is a proof of a lemma, theorem, or corollary, it ends with a  $\square$ .

Another idiosyncrasy: if a line in multiline displayed equation ends with “=”, then the next line begins with “=”. If it ends with “+”, then the next line

begin with “+”, and if it ends with “-” the next line begins with “+” as well. I’ll try to avoid ending a line with “+” or “-”, especially with “-”, but if it happens you are warned. A silly example is

$$(x + y)^2 - \tag{1}$$

$$+ 4x = \tag{2}$$

$$= x^2 + y^2 - 2x = \tag{3}$$

$$= (x - y)^2 \geq \tag{4}$$

$$\geq 0. \tag{5}$$

Just as an aside: if I refer to something that has been mentioned “above” I mean something that comes earlier in the book and “below” refers to anything that comes later. This always confuses me, so I had to write it down.

The dilemma of whether to use “we” or “I” throughout the book is solved in the usual way. If I feel that a result is the work of a group (me, my co-workers, and the giants on whose shoulders we stand) then I use “we”. If it’s an individual decision, or something personal, then I use “I”. The default is “we”, as it always should be in scientific writing.

Most of the individual chapters also have some of the necessary mathematical background material, both notation and results, sometimes with specific elaborations that seem useful for the book. Sometimes this background material is quite extensive. Examples are splines, majorization, unweighting, monotone regression, and the basic Zangwill and Ostrowski fixed point theorems we need for convergence analysis of our algorithms.

There is an appendix ?? with code, and an appendix ?? with data sets. These contain brief descriptions and links to the supplementary materials directories, which contain the actual code and data.

Something about code and R/C

I will use this note to thank Rstudio, in particular J.J. Allaire and Yihui Xi, for their contributions to the R universe, and for their promotion of open source software and open access publications. Not too long ago I was an ardent LaTeX user, firmly convinced I would never use anything else again in my lifetime. In the same way that I was convinced before I would never use anything besides, in that order, FORTRAN, PL/I, APL, and (X)Lisp.

And PHP/Apache/MySQL. But I lived too long. And then, in my dotage, lo and behold, R, Rstudio, (R)Markdown, bookdown, blogdown, Git, Github, Netlify came along.



Figure 1: Forrest Young, Bepi Pinner, Jean-Marie Bouroche, Yoshio Takane, Jan de Leeuw at La Jolla, August 1975



# Preface

This book is definitely *not* an impartial and balanced review of all of multi-dimensional scaling (MDS) theory and history. It emphasizes computation, and the mathematics needed for computation. In addition, it is a summary of over 50 years of MDS work by me, either solo or together with my many excellent current or former co-workers and co-authors. It is heavily biased in favor of the smacof formulation of MDS (De Leeuw (1977), De Leeuw and Heiser (1977), De Leeuw and Mair (2009)), and the corresponding majorization (or MM) algorithms. And, moreover, I am shamelessly squeezing in as many references to my published and unpublished work as possible, with links to the corresponding pdf's if they are available. Thus this book is also a jumpstation into my bibliography.

I have not organized the book along historical lines because most of the early techniques and results have been either drastically improved or completely abandoned. Nevertheless, some personal historical perspective may be useful. I will put most of it in this preface, so uninterested readers can easily skip it.

I got involved in MDS in 1968 when John van de Geer returned from a visit to Clyde Coombs in Michigan and started the Department of Data Theory in the Division of Social Sciences at Leiden University. I was John's first hire, although I was still a graduate student at the time.

Remember that Clyde Coombs was running the Michigan Mathematical Psychology Program, and he had just published his remarkable book "A Theory of Data" (Coombs (1964)). The name of the new department in Leiden was taken from the title of that book, and Coombs was one of the first visitors to give a guest lecture there.

This is maybe the place to clear up some possible misunderstandings about the name "Data Theory". Coombs was mainly interested in a taxonomy of

data types, and in pointing out that “data” were not limited to a table or data-frame of objects by variables. In addition, there were also similarity ratings, paired comparisons, and unfolding data. Coombs also emphasized that data were often non-metric, i.e. ordinal or categorical, and that it was possible to analyze these ordinal or categorical relationships directly, without first constructing numerical scales to which classical techniques could be applied. One of the new techniques discussed in Coombs (1964) was an ordinal form of MDS, in which not only the data but also the representation of the data in Euclidean space were non-metric.

John van de Geer had just published Van de Geer (1967). In that book, and in the subsequent book Van de Geer (1971), he developed his unique geometric approach to multivariate analysis. Relationship between variables, and between variables and individuals, were not just discussed using matrix algebra, but were also visualized in diagrams. This was related to the geometric representations in Coombs’ Theory of Data, but it concentrated on numerical data in the form of rectangular matrices of objects by variables.

Looking back it is easy to see that both Van de Geer and Coombs influenced my approach to data analysis. I inherited the emphasis on non-metric data and on visualization. But, from the beginning, I interpreted “Data Theory” as “Data Analysis”, with my emphasis shifting to techniques, loss functions, implementations, algorithms, optimization, computing, and programming. This is of interest because in 2020 my former Department of Statistics at UCLA, together with the Department of Mathematics, started a bachelor’s program in Data Theory, in which “Emphasis is placed on the development and theoretical support of a statistical model or algorithmic approach. Alternatively, students may undertake research on the foundations of data science, studying advanced topics and writing a senior thesis.” This sounds like a nice hybrid of Data Theory and Data Analysis, with a dash of computer science mixed in.

Computing and optimization were in the air in 1968, not so much because of Coombs, but mainly because of Roger Shepard, Joe Kruskal, and Doug Carroll at Bell Labs in Murray Hill. John’s other student Eddie Roskam and I were fascinated by getting numerical representations from ordinal data by minimizing explicit least squares loss functions. Eddie wrote his dissertation in 1968 (Roskam (1968)). In 1973 I went to Bell Labs for a year, and Eddie went to Michigan around the same time to work with Jim Lingoes, resulting



in Lingoes and Roskam (1973).

My first semi-publication was De Leeuw (1968), quickly followed by a long sequence of other, admittedly rambling, internal reports. Despite this very informal form of publication the sheer volume of them got the attention of Joe Kruskal and Doug Carroll, and I was invited to spend the academic year 1973-1974 at Bell Laboratories. That visit somewhat modified my cavalier approach to publication, but I did not become half-serious in that respect until meeting with Forrest Young and Yoshio Takane at the August 1975 US-Japan seminar on MDS in La Jolla. Together we used the alternating least squares approach to algorithm construction that I had developed since 1968 into a quite formidable five-year publication machine, with at its zenith Takane, Young, and De Leeuw (1977).

In La Jolla I gave the first presentation of the majorization method for MDS, later known as smacof, with the first formal convergence proof. The canonical account of smacof was published in a conference paper (De Leeuw (1977)). Again I did not bother to get the results into a journal or into some other more effective form of publication. The basic theory for what became known as smacof was also presented around the same time in another book chapter De Leeuw and Heiser (1977).

In 1978 I was invited to the Fifth International Symposium on Multivariate Analysis in Pittsburgh to present what became De Leeuw and Heiser (1980). There I met Nan Laird, one of the authors of the basic paper on the EM algorithm (Dempster, Laird, and Rubin (1977)). I remember enthusiastically telling her on the conference bus that EM and smacof were both special case of the general majorization approach to algorithm construction, which was consequently born around the same time. But that is a story for a companion volume, which currently only exists in a very preliminary stage (<https://github.com/deleeuw/bras>).

My 1973 PhD thesis (De Leeuw (1973), reprinted as De Leeuw (1984)) was actually my second attempt at a dissertation. I had to get a PhD, any PhD, before going to Bell Labs, because of the difference between the Dutch and American academic title and reward systems. I started writing a dissertation on MDS, in the spirit of what later became De Leeuw and Heiser (1982). But halfway through I lost interest and got impatient, and I decided to switch to nonlinear multivariate analysis. This second attempt did produced a finished dissertation (De Leeuw (1973)), which grew over time, with the

help of multitudes, into Gifi (1990). But that again is a different history, which I will tell some other time in yet another companion volume (<https://github.com/deleeuw/gifi>). For a long time I did not do much work on MDS, until the arrival of Patrick Mair and the R language led to a resurgence of my interest, and ultimately to De Leeuw and Mair (2009) and (?).

I consider this MDS book to be a summary and extension of the basic papers De Leeuw (1977), De Leeuw and Heiser (1977), De Leeuw and Heiser (1980), De Leeuw and Heiser (1982), and De Leeuw (1988) (published version of (?)), all written 30-40 years ago. Footprints in the sands of time. It can also be seen as an elaboration of the more mathematical and computational sections of the excellent and comprehensive textbook of Borg and Groenen (2005). That book has much more information about the origins, the data, and the applications of MDS, as well as on the interpretation of MDS solutions. In this book I concentrate almost exclusively on the mathematical, computational, and programming aspects of MDS.

For those who cannot get enough of me, there is a data base of my published and unpublished reports and papers since 1965, with links to pdf's, at <https://jansweb.netlify.app/publication/>.

There are many, many people I have to thank for my scientific education. Sixty years is a long time, and consequently many excellent teachers and researchers have crossed my path. I will gratefully mention the academics who had a major influence on my work and who are not with us any more, since I will join them in the not too distant future: Louis Guttman (died 1987), Clyde Coombs (died 1988), Warren Torgerson (died 1999), Forrest Young (died 2006), John van de Geer (died 2008), Joe Kruskal (died 2010), Doug Carroll (died 2011), and Rod McDonald (died 2012).

# Notation and Reserved Symbols

intro

## 0.1 Spaces

- $\mathbb{R}^n$  is the space of all real vectors, i.e. all  $n$ -element tuples of real numbers. Typical elements of  $\mathbb{R}^n$  are  $x, y, z$ . The element of  $x$  in position  $i$  is  $x_i$ . Defining a vector by its elements is done with  $x = \{x_i\}$ .
- $\mathbb{R}^n$  is equipped with the inner product  $\langle x, y \rangle = x'y = \sum_{i=1}^n x_i y_i$  and the norm  $\|x\| = \sqrt{x'x}$ .
- The canonical basis for  $\mathbb{R}^n$  is the  $n$ -tuple  $(e_1, \dots, e_n)$ , where  $e_i$  has element  $i$  equal to  $+1$  and all other elements equal to zero. Thus  $\|e_i\| = 1$  and  $\langle e_i, e_j \rangle = \delta^{ij}$ , with  $\delta^{ij}$  the Kronecker delta (equal to one if  $i = j$  and zero otherwise). Note that  $x_i = \langle e_i, x \rangle$ .
- $\mathbb{R}$  is the real line and  $\mathbb{R}_+$  is the half line of non-negative numbers.
- $\mathbb{R}^{n \times m}$  is the space of all  $n \times m$  real matrices. Typical elements of  $\mathbb{R}^{n \times m}$  are  $A, B, C$ . The element of  $A$  in row  $i$  and column  $j$  is  $a_{ij}$ . Defining a matrix by its elements is done with  $A = \{a_{ij}\}$ .
- $\mathbb{R}^{n \times m}$  is equipped with the inner product  $\langle A, B \rangle = \text{tr} A'B = \sum_{i=1}^n \sum_{j=1}^m a_{ij} b_{ij}$  and the norm  $\|A\| = \sqrt{\text{tr} A'A}$ .

- The canonical basis for  $\mathbb{R}^{n \times m}$  is the  $nm$ -tuple  $(E_{11}, \dots, E_{nm})$ , where  $E_{ij}$  has element  $(i, j)$  equal to +1 and all other elements equal to zero. Thus  $\|E_{ij}\| = 1$  and  $\langle E_{ij}, E_{kl} \rangle = \delta^{ik} \delta^{jl}$ .

vec and  $\text{vec}^{-1}$

## 0.2 Matrices

- $a_{i\bullet}$  is row  $i$  of matrix  $A$ ,  $a_{\bullet j}$  is column  $j$ .
- $a_{i\star}$  is the sum of row  $i$  of matrix  $A$ ,  $a_{\star j}$  is the sum of column  $j$ .
- $A'$  is the transpose of  $A$ , and  $\text{diag}(A)$  is the diagonal matrix with the diagonal elements of  $A$ . The inverse of a square matrix  $A$  is  $A^{-1}$ , the Moore-Penrose generalized inverse of any matrix  $A$  is  $A^+$ .
- If  $A$  and  $B$  are two  $n \times m$  matrices then their Hadamard (or elementwise) product  $C = A \times B$  has elements  $c_{ij} = a_{ij} b_{ij}$ . The Hadamard quotient is  $C = A/B$ , with elements  $c_{ij} = a_{ij}/b_{ij}$ . The Hadamard power is  $A^{(k)} = A^{(p-1)} \times A$ .
- DC matrices. Centering matrix.  $J_n = I_n - n^{-1}E_n$ . We do not use the subscripts if the order is obvious from the context.

## 0.3 Functions

- $f, g, h, \dots$  are used for functions or mappings.  $f : X \rightarrow Y$  says that  $f$  maps  $X$  into  $Y$ .
- $\sigma$  is used for all real-valued least squares loss functions.

## 0.4 MDS

- $\Delta = \{\delta_{ij\dots}\}$  is a matrix or array of dissimilarities.

- $\langle \mathbb{X}, d \rangle$  is a metric space, with  $d : \mathcal{X} \otimes \mathcal{X} \rightarrow \mathbb{R}_+$  the distance function. If  $X$  is an ordered  $n$ -tuple  $(x_1, \dots, x_n)$  of elements of  $\mathcal{X}$  then  $D(X)$  is  $\{d(x_i, x_j)\}$ , the elements of which we also write as  $d_{ij}(X)$ .
- Summation over the elements of vector  $x \in \mathbb{R}^n$  is  $\sum_{i=1}^n x_i$ . Summation over the elements of matrix  $A \in \mathbb{R}^{n \times m}$  is  $\sum_{i=1}^n \sum_{j=1}^m a_{ij}$ . Summation over the elements above the diagonal of  $A$  is  $\sum \sum_{1 \leq i < j \leq n} a_{ij}$ .
- Conditional summation is, for example,  $\sum_{i=1}^n \{x_i \mid x_i > 0\}$ .

Iteration



# Chapter 1

## Introduction

Placeholder

## 1.1 Brief History

### 1.1.1 Milestones

## 1.2 Basic MDS

### 1.2.1 Kruskal's stress

#### 1.2.1.1 Square root

#### 1.2.1.2 Weights

#### 1.2.1.3 Normalization

## 1.3 Local and Global

## 1.4 Generalizations

### 1.4.1 Non-metric MDS

### 1.4.2 fstress

### 1.4.3 Constraints

### 1.4.4 Individual Differences

### 1.4.5 Distance Asymmetry

## 1.5 Models and Techniques



## Chapter 2

# Properties of Stress

Placeholder



## 2.1 Notation

### 2.1.1 Expanding

### 2.1.2 Matrix Expressions

### 2.1.3 Coefficient Space

### 2.1.4 Our Friends CS and AM/GM

## 2.2 Global Properties

### 2.2.1 Boundedness

### 2.2.2 Invariance

### 2.2.3 Continuity

### 2.2.4 Coercivity

## 2.3 Differentiability

### 2.3.0.1 Distances

### 2.3.0.2 Rho and Stress

### 2.3.0.3 `expandStress`

### 2.3.1 Partial Derivatives

### 2.3.2 Special Expansions

#### 2.3.2.1 Infinitesimal Rotations

#### 2.3.2.2 Singularities

#### 2.3.2.3 Singularities

## 2.4 Convexity

### 2.4.1 Distances

### 2.4.2 Subdifferentials



# Chapter 3

## Stress Spaces

Placeholder



## 3.1 Configuration Space

### 3.1.1 Zero Distance Subspaces

## 3.2 Spherical Space

## 3.3 Distance Space

## 3.4 Squared Distance Space

## 3.5 Gramian Space

## 3.6 Pictures of Stress

### 3.6.1 Coefficient Space

### 3.6.2 Global Perspective

### 3.6.3 Global Contour

### 3.6.4 Stationary Points

#### 3.6.4.1 First Minimum

### 3.6.5 Second Minimum

### 3.6.6 Third Minimum

### 3.6.7 First Saddle Point

### 3.6.8 Second Saddle Point

## 3.7 Another Look

## 3.8 A Final Look

## 3.9 Discussion

## 3.10 Coordinates





# Chapter 4

## Classical Multidimensional Scaling

Placeholder

### 4.1 Algebra

#### 4.1.1 Torgerson Transform

#### 4.1.2 Schoenberg's Theorem

### 4.2 Approximation

#### 4.2.1 Low Rank Approximation of the Torgerson Transform

#### 4.2.2 Minimization of Strain

#### 4.2.3 Approximation from Below



# Chapter 5

## Minimization of Basic Stress

Placeholder



## 5.1 Gradient Methods

### 5.1.1 Step Size

#### 5.1.1.1 Kruskal Step Size

#### 5.1.1.2 Guttman Step Size

#### 5.1.1.3 Cauchy Step Size

#### 5.1.1.4 Majorization Step Size

## 5.2 Initial Configurations

## 5.3 On MM Algorithms

## 5.4 Smacof Algorithm

### 5.4.1 Guttman Transform

### 5.4.2 Subdifferentials

### 5.4.3 Derivative

### 5.4.4 Global Convergence

#### 5.4.4.1 From the CS Inequality

#### 5.4.4.2 From Majorization

#### 5.4.4.3 From Ratio of Norms

### 5.4.5 Component Rotated Smacof

### 5.4.6 Local Convergence

#### 5.4.6.1 Cross Product Iterations

#### 5.4.6.2 Rotation to PC

### 5.4.7 Data Asymmetry

### 5.4.8 Replications



# Chapter 6

## Acceleration of Convergence

Placeholder

### 6.1 Simple Acceleration

### 6.2 One-Parameter Methods

### 6.3 SQUAREM

### 6.4 Vector Extrapolation Methods





# Chapter 7

## Nonmetric MDS

Placeholder

## **7.1 Generalities**

### **7.1.1 Kruskal's Stress**

## **7.2 Single and Double Phase**

### **7.2.1 Double Phase**

### **7.2.2 Single Phase**

## **7.3 Affine NMDS**

## **7.4 Conic NMDS**

### **7.4.1 Normalization**

### **7.4.2 Normalized Cone Regression**

### **7.4.3 Hard Squeeze and Soft Squeeze**

### **7.4.4 Inner Iterations**

### **7.4.5 Stress-1 and Stress-2**

# Chapter 8

## Interval MDS

Placeholder

## 8.1 The Additive Constant

### 8.1.1 Early

### 8.1.2 Cooper

## 8.2 Algebra

### 8.2.1 Existence

### 8.2.2 Solution

### 8.2.3 Examples

#### 8.2.3.1 Small Example

#### 8.2.3.2 De Gruijter Example

#### 8.2.3.3 Ekman Example

### 8.2.4 A Variation

## 8.3 Interval smacof

### 8.3.1 Example

# Chapter 9

## Polynomial MDS

Placeholder

## 9.1 Introduction

## 9.2 Fitting Polynomials

## 9.3 Positive and Convex, Monotone Polynomials

### 9.3.1 Introduction

### 9.3.2 A QP Algorithm

#### 9.3.2.1 Example 1: Simple Monotone Regression

#### 9.3.2.2 Example 2: Monotone Regression with Ties

#### 9.3.2.3 Example 3: Weighted Rounding

#### 9.3.2.4 Example 4: Monotone Polynomials

### 9.3.3 Examples

# Chapter 10

## Ordinal MDS

Placeholder

## 10.1 Monotone Regression

### 10.1.1 Simple Monotone Regression

### 10.1.2 Weighted Monotone Regression

### 10.1.3 Normalized Cone Regression

### 10.1.4 Iterative MR

## 10.2 Alternating Least Squares

## 10.3 Kruskal's Approach

### 10.3.1 Kruskal's Stress

### 10.3.2 Stress1 and Stress2

## 10.4 Guttman's Approach

### 10.4.0.1 Rank Images

### 10.4.0.2 Single and Double Phase

### 10.4.0.3 Hard and Soft Squeeze

### 10.4.1 Smoothness of Ordinal Loss Functions

## 10.5 Scaling with Distance Bounds

## 10.6 Bounds on Stress



# Chapter 11

## Splinal MDS

Placeholder

## 11.1 Splines

### 11.1.1 B-splines

#### 11.1.1.1 Boundaries

#### 11.1.1.2 Normalization

#### 11.1.1.3 Recursion

#### 11.1.1.4 Illustrations

### 11.1.2 I-splines

#### 11.1.2.1 Increasing Coefficients

#### 11.1.2.2 Increasing Values

### 11.1.3 Time Series Example

#### 11.1.3.1 B-splines

#### 11.1.3.2 I-splines

#### 11.1.3.3 B-Splines with monotone weights

#### 11.1.3.4 B-Splines with monotone values

### 11.1.4 Local positivity, monotonicity, convexity

# Chapter 12

## Unidimensional Scaling

Placeholder

## 12.1 An example

### 12.1.1 Perspective

### 12.1.2 Contour

## 12.2 Order Formulation

## 12.3 Permutation Formulation

## 12.4 Sign Matrix Formulation

## 12.5 Algorithms for UMDS

### 12.5.1 SMACOF

### 12.5.2 SMACOF (smoothed)

### 12.5.3 Branch-and-Bound

### 12.5.4 Dynamic Programming

### 12.5.5 Simulated Annealing

### 12.5.6 Penalty Methods

# Chapter 13

## Full-dimensional Scaling

Placeholder

**13.1 Convexity**

**13.2 Optimality**

**13.3 Iteration**

**13.4 Cross Product Space**

**13.5 Full-dimensional Scaling**

**13.6 Ekman example**



# Chapter 14

## Unfolding

Placeholder

### 14.1 Algebra

#### 14.1.1 One-dimensional

### 14.2 Classical Unfolding

### 14.3 Nonmetric Unfolding

#### 14.3.1 Degenerate Solutions

##### 14.3.1.1 Which Stress

##### 14.3.1.2 l'Hôpital's Rule

##### 14.3.1.3 Penalizing

##### 14.3.1.4 Restricting Regression





## Chapter 15

# Constrained Multidimensional Scaling

Placeholder



## **15.1 Primal-Dual (note: the base partitioning has dual aspects)**

### **15.2 Basic Partitioning**

### **15.3 Unweighting**

### **15.4 Constraints on the Distances**

#### **15.4.1 Rectangles**

### **15.5 Linear Constraints**

#### **15.5.1 Uniqueness**

#### **15.5.2 Combinations**

#### **15.5.3 Step Size**

#### **15.5.4 Single Design Matrix**

#### **15.5.5 Multiple Design Matrices**

### **15.6 Circular MDS**

#### **15.6.1 Some History**

#### **15.6.2 Primal Methods**

#### **15.6.3 Dual Methods**

### **15.7 Elliptical MDS**

#### **15.7.1 Primal**

#### **15.7.2 Dual**

### **15.8 Distance Bounds**



# Chapter 16

## Individual Differences

Placeholder

## **16.1 Stress Majorization**

## **16.2 Types of Constraints**

### **16.2.1 Unconstrained**

### **16.2.2 Replications**

### **16.2.3 INDSCAL/PARAFAC**

### **16.2.4 IDIOSCAL/TUCKALS**

### **16.2.5 PARAFAC2**

### **16.2.6 Factor Models**

## **16.3 Nonmetric Individual Differences**

### **16.3.1 Conditionality**

# Chapter 17

## Asymmetry in MDS

### 17.1 Conditional Rankings

Two sets (= unfolding), one set (much tighter) solution Young\_75

### 17.2 Confusion Matrices

Choice theory

### 17.3 The Slide Vector

### 17.4 DEDICOM

### 17.5 Constantine-Gower





# Chapter 18

## Nominal MDS

Placeholder

18.1 Binary Dissimilarities

18.2 Indicator matrix

18.3 Unfolding Indicator Matrices

18.4 Linear Separation

18.5 Circular Separation

18.6 Convex Hull Scaling

18.7 Voronoi Scaling

18.8 Multidimensional Scalogram Analysis



# Chapter 19

## Nonmonotonic MDS

Placeholder

### 19.1 Filler



# Chapter 20

## Compound Objects

### 20.1 Filler

A compound object is a set of  $m > 1$  objects. This chapter treats the analysis of dissimilarities between compound objects, again as far as they fit into the smacof framework. Thus we do not, for example, look at the pairwise dissimilarities within a triad of objects, but at the dissimilarities of the triads themselves.



# Chapter 21

## Stress and strain

Placeholder

## **21.1 sstress**

### **21.1.1 sstress and stress**

### **21.1.2 Decomposition**

### **21.1.3 Full-dimensional sstress**

### **21.1.4 Minimizing sstress**

#### **21.1.4.1 ALSCAL**

#### **21.1.4.2 Majorization**

#### **21.1.4.3 SOS**

#### **21.1.4.4 Duality**

#### **21.1.4.5 ALS Unfolding**

### **21.1.5 Bounds for sstress**

## **21.2 strain**

### **21.2.1 Unweighted**

### **21.2.2 Bailey-Gower**

### **21.2.3 Using Additivity**



# Chapter 22

## fstress and rstress

Placeholder

### 22.1 fstress

#### 22.1.1 Use of Weights

#### 22.1.2 Convexity

### 22.2 rStress

#### 22.2.1 Using Weights

#### 22.2.2 Minimizing rstress

### 22.3 mstress

### 22.4 astress

### 22.5 pstress



# Chapter 23

## Alternative Least Squares Loss

Kruskal/Carroll in Krishnaiah

### 23.1 Sammon's MDS

### 23.2 Kamade-Kawai Spring

### 23.3 McGee's Work

### 23.4 Shepard's Nonmetric MDS

### 23.5 Guttman's Nonmetric MDS

### 23.6 Positive Orthant Nonmetric MDS

!! Richard Johnson 1973

!! Guttman Absolute Value

!! Hartmann

## 23.7 Role Reversal

Kruskal, arithmetic with dissimilarities

$$\sigma(X) = \sum_{1 \leq i < j \leq n} \sum w_{ij} (\delta_{ij} - P_r(d_{ij}(X)))^2$$

## Chapter 24

# Inverse Multidimensional Scaling

Placeholder



## 24.1 Basic IMDS

## 24.2 Non-negative Dissimilarities

## 24.3 Zero Weights and/or Distances

## 24.4 Examples

### 24.4.1 First Example

### 24.4.2 Second Example

### 24.4.3 Third Example

### 24.4.4 Fourth Example

## 24.5 MDS Sensitivity

## 24.6 Second Order Inverse MDS

## 24.7 Inverse FDS

## 24.8 Multiple Solutions

## 24.9 Minimizing iStress

## 24.10 Order three

## 24.11 Order Four

## 24.12 Sensitivity





# Chapter 25

## Stability of MDS Solutions

Placeholder

### 25.1 Null Distribution

### 25.2 Pseudo-confidence Ellipsoids

### 25.3 A Pseudo-Bootstrap

### 25.4 How Large is My Stress ?



# Chapter 26

## In Search of Global Minima

Placeholder



## 26.1 Random Starts

### 26.1.1 Examples

#### 26.1.1.1 Ekman

#### 26.1.1.2 De Gruijter

## 26.2 Tunneling, Filling, Annealing, etc.

## 26.3 Cutting Planes

### 26.3.1 On the Circle

### 26.3.2 Cauchy Step Size

### 26.3.3 Balls

#### 26.3.3.1 Outer Approximation

## 26.4 Distance Smoothing

## 26.5 Penalizing Dimensions

### 26.5.1 Local Minima

### 26.5.2 Algorithm

### 26.5.3 Examples

#### 26.5.3.1 Chi Squares

#### 26.5.3.2 Regular Simplex

#### 26.5.3.3 Intelligence

#### 26.5.3.4 Countries

#### 26.5.3.5 Dutch Political Parties

#### 26.5.3.6 Ekman

#### 26.5.3.7 Morse in Two



# Chapter 27

## Software

In actual computer output using the scaling in formula (??) and (??) has some disadvantages. There are, say,  $M$  non-zero weights. The summation in #ref(eq:stressall) is really over  $M$  terms only. If  $n$  is at all large the scaled dissimilarities, and consequently the distances and the configuration, will become very small. Thus, in actual computation, or at least in the computer output, we scale our dissimilarities as  $\frac{1}{2} \sum \sum_{1 \leq j < i \leq n} w_{ij} \delta_{ij}^2 = M$ . So, we scale our dissimilarities to one in formulas and to  $M$  in computations. Thus the computed stress will b

rcode

ccode

lib

Placeholder





## **.1 R Code**

**.1.1 utilities.R**

**.1.2 common/indexing.r**

**.1.3 common/io.r**

**.1.4 common/linear.r**

**.1.5 common/nextPC.r**

**.1.6 common/smacof.r**

**.1.7 properties.R**

**.1.8 expandOneDim.R**

**.1.9 pictures.R**

**.1.10 classical.R**

**.1.11 minimization.R**

**.1.12 full.R**

**.1.13 unfolding.R**

**.1.14 constrained.R**

**.1.15 nominal.R**

**.1.16 sstress.R**

**.1.17 inverse.R**

**.1.18 global.R**

**.1.19 mathadd.R**

## **.2 C code**

**.2.1 deboor.c**

- Coombs, C. H. 1964. *A Theory of Data*. Wiley.
- De Leeuw, J. 1968. “Nonmetric Multidimensional Scaling.” Research Note 010-68. Department of Data Theory FSW/RUL.
- . 1973. “Canonical Analysis of Categorical Data.” PhD thesis, University of Leiden, The Netherlands.
- . 1977. “Applications of Convex Analysis to Multidimensional Scaling.” In *Recent Developments in Statistics*, edited by J. R. Barra, F. Brodeau, G. Romier, and B. Van Cutsem, 133–45. Amsterdam, The Netherlands: North Holland Publishing Company.
- . 1984. *Canonical Analysis of Categorical Data*. Leiden, The Netherlands: DSWO Press.
- . 1988. “Convergence of the Majorization Method for Multidimensional Scaling.” *Journal of Classification* 5: 163–80.
- De Leeuw, J., and W. J. Heiser. 1977. “Convergence of Correction Matrix Algorithms for Multidimensional Scaling.” In *Geometric Representations of Relational Data*, edited by J. C. Lingoes, 735–53. Ann Arbor, Michigan: Mathesis Press.
- . 1980. “Multidimensional Scaling with Restrictions on the Configuration.” In *Multivariate Analysis, Volume V*, edited by P. R. Krishnaiah, 501–22. Amsterdam, The Netherlands: North Holland Publishing Company.
- . 1982. “Theory of Multidimensional Scaling.” In *Handbook of Statistics, Volume II*, edited by P. R. Krishnaiah and L. Kanal. Amsterdam, The Netherlands: North Holland Publishing Company.
- 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>.
- Dempster, A. P., N. M. Laird, and D. B. Rubin. 1977. “Maximum Likelihood for Incomplete Data via the EM Algorithm.” *Journal of the Royal Statistical Society* B39: 1–38.
- Gifi, A. 1990. *Nonlinear Multivariate Analysis*. New York, N.Y.: Wiley.
- Lingoes, J. C., and E. E. Roskam. 1973. “A Mathematical and Empirical Analysis of Two Multidimensional Scaling Algorithms.” *Psychometrika* 38: Monograph Supplement.
- Roskam, E. E. 1968. “Metric Analysis of Ordinal Data in Psychology.” PhD thesis, University of Leiden.
- Takane, Y., F. W. Young, and J. De Leeuw. 1977. “Nonmetric Individual Differences in Multidimensional Scaling: An Alternating Least Squares

Method with Optimal Scaling Features.” *Psychometrika* 42: 7–67.

Van de Geer, J. P. 1967. *Inleiding in de Multivariate Analyse*. Van Loghum Slaterus.

———. 1971. *Introduction to Multivariate Analysis for the Social Sciences*. San Francisco, CA: Freeman.