MULTIDIMENSIONAL DATA ANALYSIS

edited by

Jan De Leeuw
Willem Heiser
Jacqueline Meulman
Frank Critchley

Pembroke College

DSWO PRESS
MULTIDIMENSIONAL
DATA ANALYSIS

Proceedings of a workshop, Pembroke College,
Cambridge University, England, June 30 - July 2 1985

directed by

Jan de Leeuw
Willem Heiser
Jacqueline Meulman

Department of Data Theory
University of Leiden

Frank Critchley

Department of Statistics
University of Warwick

1986 DSWO Press, Leiden
ISBN 90-6695-015-3
SISO 517.2 UDC 519.2
Subject heading: multivariate analysis, multidimensional data analysis.

© 1986 DSWO Press, Leiden
All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without prior permission of the publisher.

Cover drawing by Courtesy of Pembroke College, Cambridge University
Printed by 'Reprodienst, Subfaculteit Psychologie, Rijksuniversiteit Leiden’ and by printing office 'Karstens drukkers bv', Leiden

ISBN 90-6695-015-3
CONTENTS

Editors' Introduction 1
Workshop Programme 5

THEME 1. ALGEBRAIC ASPECTS OF
DISTANCE-BASED METHODS 9

J. C. Gower, Euclidean distance matrices 11
B. Fichet, Distances and Euclidean distances for presence-absence characters
and their application to factor analysis 23
W. J. Heiser, Distances and their approximation 47
F. Critchley, Some observations on distance matrices 53
J. C. Gower, Reply to discussants 61

THEME 2. ALGEBRAIC ASPECTS OF SCALAR PRODUCT
AND BIPILOT ANALYSES 65

J. P. Van de Geer, Relations among k sets of variables with geometrical
representation and application to categorical variables 67
J. M. F. Ten Berge, A general solution for the Maxbet problem 81
J. P. Van de Geer, A reply to Jos Ten Berge 89
K. R. Gabriel and C. L. Odoroff, Some diagnoses of models by 3-D biplots 91
M. Greenacre, Discussion on paper by Gabriel and Odoroff 113
K. R. Gabriel, A reply to Michael Greenacre 115

THEME 3. PROBABILISTIC AND STATISTICAL ASPECTS 117

Y. Takane and T. Shibayama, Comparison of models of stimulus recognition data 119
I. W. Molenaar, Deconfusing confusion matrices 139
Y. Takane and T. Shibayama, A reply to Ivo Molenaar 147
H. Caussinus, Models and uses of principal component analysis 149
J. De Leeuw, The role of models in principal component and factor analysis 171
H. Caussinus, A reply to Jan De Leeuw 177
THEME 4. CLASSIFICATION AND SERIATION

L. Hubert and P. Arabie, Unidimensional scaling and combinatorial optimization
G. De Soete, A discussion of Hubert and Arabie's paper on unidimensional scaling
E. Diday, Orders and overlapping clusters in pyramids
P. Arabie, Comments on confounded experimental designs, pyramids, and seriation

ABSTRACTS OF CONTRIBUTED PAPERS

M. W. Browne, The Young-Householder algorithm and the least squares
multidimensional scaling of squared distances
J. D. Carroll and S. Winsberg, Maximum likelihood procedures for metric and
quasi-nonmetric fitting of an extended INDSCAL model assuming both
common and specific dimensions
I. Stoop, The contribution of merging: a comparision of two data collection
methods in MDS
G. d'Aubigny, Interpreting mappings issued from distance-based methods
E. Van der Burg and J. De Leeuw, Use of the multinomial Jackknife in
generalized canonical correlation analysis
G. Saporta, Confidence regions for Correspondence analysis
S. Nishisato, Multidimensional analysis of successive categories
M. C. Jones, An introduction to projection pursuit
F. W. Young, T. Edds, D. Kent and W. F. Kuhfeld, Visual exploratory
multivariate data analysis: a progress report
J. B. Kruskal and S. Winsberg, Some metrics for sampled smoothed functions,
based on interpolation: a revised version of some work by Besse and Ramsay
J. B. Kruskal and S. Winsberg, Easy-to-generate metrics, for use with sampled
functions, when changes are important and filtering may be desired
J. Fine, Addition of a nominal variable in principal component analysis:
propositions of metrics
A. D. Gordon and C. R. Finden, Classification of spatially-located data
G. W. Furnas, The minimum additive constant problem for distance matrices with
missing data

List of participants
MULTIDIMENSIONAL DATA ANALYSIS

There have been many recent developments in multidimensional data analysis, and to a considerable extent these developments are due to psychometricians. This book contains the proceedings of a workshop on multidimensional data analysis that was organized in Cambridge in the summer of 1985, just before the biannual European Meeting of the Psychometric Society. It covers the areas of distance models used in multidimensional scaling, of discrete and hierarchical models used in classification and cluster analysis, of inner product models used in biplot and related graphical techniques, and it also discusses various attempts to integrate these data analytical techniques with more conventional statistical approaches based on probability models. The papers that are included are either research papers, dealing with a specific topic, or review papers, which discuss the state of the art in a particular field. All main invited papers are followed by invited discussions. This 'workshop format' of the book makes it possible to see quite clearly what the main problems, approaches, and controversies are in the various fields. It is consequently an excellent introduction to those areas of multidimensional data analysis that are currently attracting much attention in psychometrics.