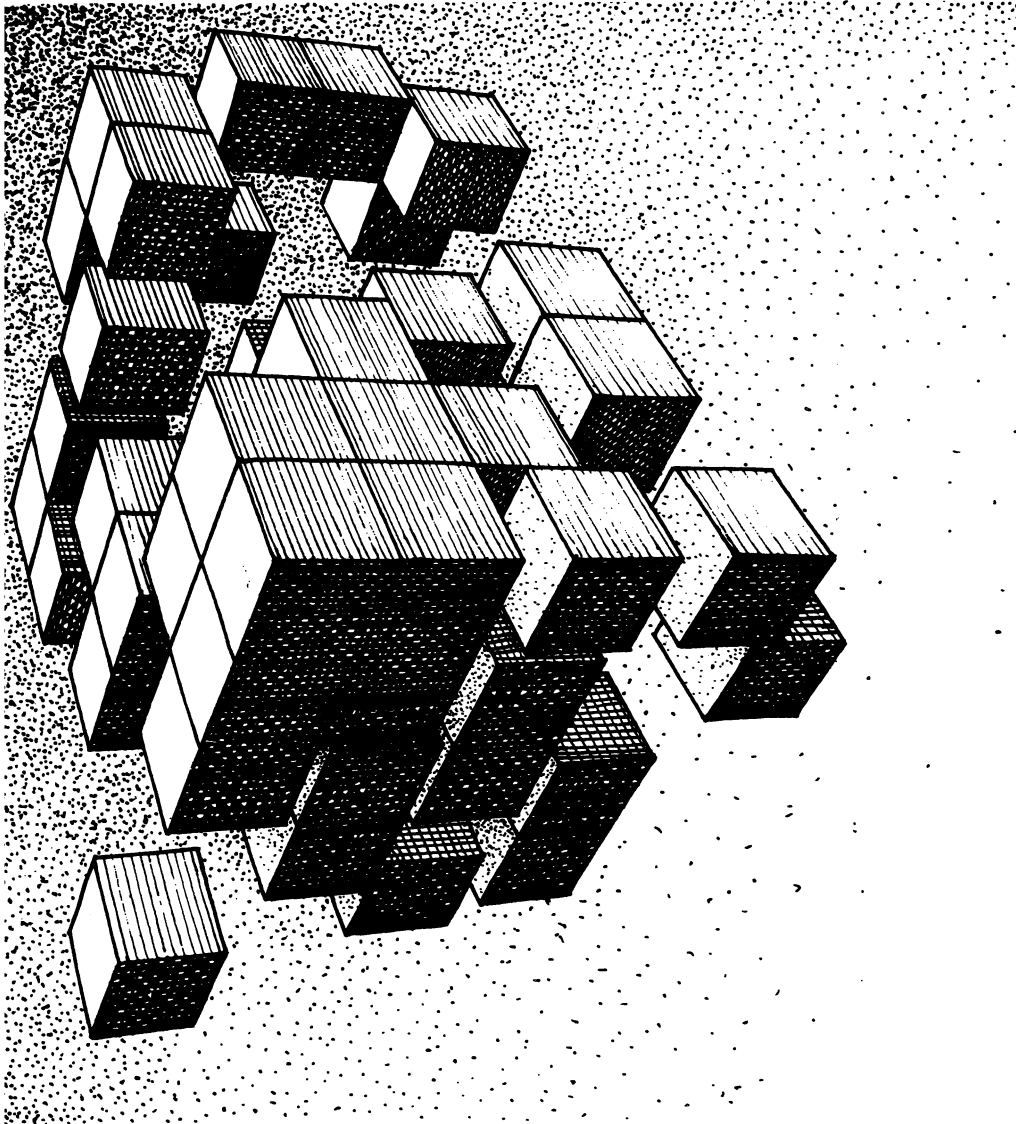


THREE-MODE
PRINCIPAL COMPONENT ANALYSIS

Pieter M. Kroonenberg



DSWO PRESS

**THREE-MODE PRINCIPAL COMPONENT ANALYSIS
THEORY AND APPLICATIONS**

M&T series volume 2

Editorial Staff:

Prof. dr. W.J. Heiser
Prof. dr. L.J.T. van der Kamp
Prof. dr. J. de Leeuw

Technical Editor:

L. Delvaux

Other publications in this series:

Jacqueline Meulman, Homogeneity analysis of incomplete data.
M&T series 1, 1982

Pieter M. Kroonenberg, Tree-mode principle component analysis.
M&T series 2, 1983, reprint 1989

Jan de Leeuw, Canonical analysis of categorical data.
M&T series 3, 1984

Ronald A. Visser, Analysis of longitudinal data in behavioural and social research
M&T series 4, 1985

John P. van de Geer, Introduction to linear multivariate data analysis.
M&T series 5, volume 1 & 2, 1986

Jacqueline Meulman, A distance approach to nonlinear multivariate
analysis.
M&T series 6, 1986

Jan de Leeuw, Willem Heiser, Jacqueline Meulman, Frank Critchley
(editors), Multidimensional data analysis.
M&T series 7, 1987

Peter G.M. van der Heijden, Correspondence analysis of longitudinal
categorical data.
M&T series 8, 1987

Jan van Rijckevorsel, The application of fuzzy coding and horseshoes in multiple
correspondence analysis.
M&T series 9, 1987

Abby Israëls, Eigenvalue techniques for qualitative data.
M&T series 10, 1987

Eeke van der Burg, Nonlinear canonical correlation and some related techniques.
M&T series 11, 1988

Kees van Montfort, Estimating in structural models with non-normal distributed
variables: some alternative approaches.
M&T series 12, 1989

Available from DSWO Press
Wassenaarseweg 52
2333 AK Leiden
The Netherlands
Tel. (071 - 273795)

THREE-MODE
PRINCIPAL COMPONENT ANALYSIS
THEORY AND APPLICATIONS

Pieter M. Kroonenberg

*Department of Data Theory
Faculty of Social Sciences
University of Leiden*

1983 DSWO Press, Leiden
Reprint 1989

CIP-DATA KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Kroonenberg, Pieter M.

Three-mode principal component analysis: theory and applications / Pieter M Kroonenberg. – Leiden: DSWO Press. – Ill. – (M&T series; vol. 2)

Ook verschenen als proefschrift Leiden. – Met index, lit. opg. – Met samenvatting in het Nederlands.

ISBN 90-6695-002-1

SISO 300.6 UDC 301.06

Trefw.: sociaal-wetenschappelijk onderzoek / statistiek / data-analyse.

© 1983 DSWO Press, Leiden, reprint 1989

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 design and Prelims Marjorie Meulman

Cover drawing Siep Kroonenberg

Printed by 'Reprodienst, Subfaculteit Psychologie, Rijksuniversiteit Leiden' and by Printing office 'Karstens drukkers bv, Leiden'

ISBN 90-6695-002-1

"...statisticians have no fields of their own from which to harvest their data. Statisticians get all their data from other fields, and from all other fields, wherever data are gathered ...

Ours is a symbiotic way of life, a marginal and hyphenated existence. We resemble the professional harvesters of wheat and grains on [the] Great Plains, who own no field of their own, but harvest field after field, in state after state, and lead a useful, rewarding, and interesting life - as we do."

Leslie Kish, *JASA*, 1978, 73, 1

PREFACE

The purpose of this book is three-fold. In the first place it is a monograph on three-mode principal component analysis. An attempt has been made to discuss virtually all issues in connection with the technique, and to collect and evaluate the literature on the subject which has appeared since its introduction in the sixties by Ledyard R. Tucker at the University of Illinois.

Secondly, this book introduces improved estimation procedures of the parameter in three-mode principal component models, and treats a number of consequences of these new procedures. Furthermore, some theoretical contributions with respect to transformations of core matrices are presented.

Thirdly, this book aims to provide a guide for social scientists and others who wish to apply three-mode analysis to their data. General issues with regard to what goes in and what comes out of a three-mode analysis are discussed, while some thirteen data sets from varying backgrounds and composition are analysed. This book also brings together almost the entire applied literature as part of the references, thus allowing researchers to compare their results with others from their own field of interest. A classification of these applied papers has been included in the Appendix.

Many people assisted in the preparation of this book. The core of this book and many technical details were conceived by Jan de Leeuw. Leo van der Kamp, Wim van der Kloot, and Charles Lewis read various parts of the book, suggested improvements, corrected errors, and helped clarify many ideas expressed here. Frits Goossens, Ineke Stoop, Jan Swaan, Albert Verbeek, Ron Visser, and Rien van IJzendoorn assisted with parts of the text. Jeanet Bus, Peter van der Heijden, Wim van der Kloot, Cor Lammers, Jan de Leeuw, and Tom van der Voort provided or suggested data sets for inclusion. Ineke Smit polished the English as far as was possible without specific knowledge of the subject matter; Siep Kroonenberg made the cover draw-

ing; Cora Jongsma expertly typed the text, and patiently included the many formulas and revisions; Piet Brouwer joined in the drudgery of checking the manuscript; Lutgart Balfoort assisted in getting this book into print.

Leiden, January 1983

CONTENTS

Preface	
Contents	
Detailed contents	
1. Preliminaries	1
2. Survey	13
Part I: THEORY	45
3. Models	47
4. Methods and algorithms	75
5. Transformations of core matrices	107
PART II: THEORY FOR APPLICATIONS	125
6. Scaling and interpretation	127
7. Residuals	169
PART III: APPLICATIONS	197
8. Standard three-mode data: Attachment study	201
9. Semantic differential data: Triple personality study	227
10. Asymmetric similarity data: ITP study	243
11. Similarities and adjective ratings: Cola study	255
12. Correlation matrices: Four ability-factor study	273
13. Multivariate longitudinal data: Hospital study	285
14. Growth curves: Learning-to-read study	313
15. Three-mode correspondence analysis: Leiden electorate study	325
Samenvatting	345
Acknowledgements	351
Appendices	353
References	359
Author index	387
Subject index	389

DETAILED CONTENTS

PREFACE
CONTENTS
DETAILED CONTENTS

1. PRELIMINARIES	1
1.1 Introduction	2
1.2 Some examples	3
<i>semantic differential data</i>	
<i>similarity data</i>	
<i>asymmetric similarity data</i>	
<i>multivariate longitudinal data</i>	
<i>three-way contingency tables</i>	
1.3 Organization of this book	5
1.4 Three-mode glossary	7
<i>basic terms</i>	
<i>methods</i>	
<i>models</i>	
<i>terms with special definitions</i>	
1.5 Notation	11
2. SURVEY	13
2.1 Introduction	14
2.2 Informal descriptions	14
<i>research questions arising from three-mode data</i>	
<i>structure: raw scores derived from idealized elements</i>	
<i>methodology: extending singular value decomposition</i>	
2.3 Formal descriptions	21
<i>Tucker3 model</i>	
<i>Tucker2 model</i>	
2.4 Interpretational aids	24
<i>joint plots</i>	
<i>component scores</i>	
<i>residuals</i>	
<i>scaling of input data</i>	
2.5 Party similarity study: design and data	26
2.6 Analysis and fit	29
<i>analyses</i>	
<i>fit</i>	
2.7 Configurations of the three modes	32

2.8	Core matrices	35
	<i>TUCKALS3 core matrix</i>	
	<i>TUCKALS2 extended core matrix</i>	
2.9	Fit of preference groups and stimuli	38
2.10	Joint plots and component scores	41
P A R T I: T H E O R Y		45
	Summary	
3.	MODELS	47
3.1	Introduction	48
3.2	Component models with three reduced modes	52
	<i>Tucker3 model</i>	
	<i>three-mode scaling</i>	
	<i>parallel factor analysis (PARAFAC1)</i>	
	<i>individual differences scaling (INDSCAL)</i>	
3.3	Component models with two reduced modes	54
	<i>Tucker2 model</i>	
	<i>individual differences in orientation scaling (IDIOSCAL)</i>	
	<i>parallel factor analysis (PARAFAC2)</i>	
	<i>canonical decomposition (CANDECOMP)</i>	
	<i>individual differences scaling (INDSCAL)</i>	
3.4	Generality of the Tucker3 model	57
3.5	Factor analysis models or covariance structure models	60
	<i>Introduction</i>	
	<i>Tucker's three-mode common factor analysis model</i>	
	<i>Snyder's unique variances model</i>	
	<i>Bloxom's reformulation of Tucker's model</i>	
	<i>the Bentler & Lee models</i>	
	<i>three-mode factor analysis as a covariance structure model</i>	
3.6	Individual differences models or covariance structure models?	66
3.7	Extensions of three-mode principal component models	68
	<i>missing data</i>	
	<i>extensions to other measurement characteristics</i>	
	<i>external analysis</i>	
	<i>restrictions on configurations</i>	
3.8	Three-mode causal modelling	71
3.9	n-mode extensions	73
4.	METHODS AND ALGORITHMS	75
4.1	Introduction	76
4.2	Tucker's methods for three-mode principal component analysis	76
4.3	Least squares solutions for the Tucker3 model	79
	<i>loss functions</i>	
	<i>existence of a best approximate solution</i>	

	<i>partitioning of total sum of squares</i>	
	<i>nature of approximate solution</i>	
	<i>nature of exact solution</i>	
4.4	Alternating least squares algorithm for Tucker3 model	86
	<i>introduction</i>	
	<i>alternating least squares approach</i>	
	<i>simultaneous iteration method</i>	
	<i>TUCKALS3 algorithm</i>	
	<i>convergence of TUCKALS3 algorithm</i>	
4.5	Nesting of components and initialization	92
	<i>nested solutions?</i>	
	<i>initialization</i>	
	<i>upper bounds for SS(Fit)</i>	
4.6	Alternating least squares algorithm for Tucker2 model	95
	<i>TUCKALS2 algorithm</i>	
4.7	Computational accuracy and propagation of errors	96
	<i>Hilbert cubes and replicated Hilbert matrices</i>	
	<i>propagation of errors in similarity judgements</i>	
4.8	Conclusion	102
	Appendix 4.1 Proof that $SS(Tot) = SS(Fit) + SS(Res)$	
	Appendix 4.2 Bounds for SS(Fit)	
5.	TRANSFORMATIONS OF CORE MATRICES	107
5.1	Introduction	108
5.2	Orthonormal transformations	110
	<i>problem and solution</i>	
	<i>algorithm</i>	
5.3	Non-singular transformations	112
	<i>problem and solution</i>	
	<i>algorithm</i>	
5.4	Comparison of transformation procedures	116
5.5	Illustrations of transformations	119
	<i>Four ability-factor study</i>	
	<i>Perceived reality study</i>	
5.6	Concluding remarks	123
PART II: THEORY FOR APPLICATIONS		125
	Summary	
6.	SCALING AND INTERPRETATION	127
6.1	Introduction	128
6.2	Input scaling: general considerations	129
	<i>types of scaling</i>	
	<i>selecting a type of scaling</i>	
	<i>types of three-mode data</i>	
6.3	Input scaling: arbitrary and incomparable means and variances	133

	<i>arbitrary means and variances</i>	
	<i>incomparable means and variances</i>	
6.4	Input centring: interpretable means <i>two-mode data</i> <i>three-mode data</i>	135
6.5	Input centring: types, consequences, recommendations <i>types of centring</i> <i>some consequences of centring</i> <i>recommendations</i>	142
6.6	Input standardization: comparable variances	149
6.7	Interpretation: general issues	151
6.8	Interpretation: components <i>components as latent elements</i> <i>scaling to the size of component weights</i> <i>scaling according to Bartussek</i> <i>rotation of components</i>	154
6.9	Interpretation: core matrices <i>explained variation</i> <i>three-mode interactions</i> <i>scores of idealized elements</i> <i>direction cosines</i>	157
6.10	Interpretation: joint plots and component scores <i>joint plots</i> <i>component scores</i> <i>mixed-mode matrices</i>	164
7.	RESIDUALS	169
7.1	Introduction	170
7.2	Informal inference and goals of residual analysis	170
7.3	Procedures for analysing residuals <i>principal component residuals from two-mode data</i> <i>least squares residuals from two-mode data</i> <i>three-mode residuals</i>	172
7.4	Scheme for the analysis of three-mode residuals	177
7.5	An illustrative data set: Perceived reality study	180
7.6	Residual analysis for Perceived reality study <i>distributions of total and residual sums of squares</i> <i>analysis-of-variance decomposition of sums of squares</i> <i>sums-of-squares plots</i> <i>individual residuals (unstructured approach)</i>	184
7.7	Three-mode analysis and three-way ANOVA decomposition	194
PART III: APPLICATIONS		197
	Summary	
8.	STANDARD THREE-MODE DATA: ATTACHMENT STUDY	201
8.1	Design and data description	202
8.2	Analysis and fit	206

8.3	Configurations of the three modes	207
	<i>episodes</i>	
	<i>interactive scales</i>	
	<i>children</i>	
8.4	Interpretation of the core matrices	213
	<i>explained variation</i>	
	<i>three-mode interactions</i>	
	<i>extended core matrix</i>	
8.5	Joint plots	219
8.6	Fit of the scales, episodes, and children	221
8.7	Discussion	224
9.	SEMANTIC DIFFERENTIAL DATA: TRIPLE PERSONALITY STUDY	227
9.1	Introduction	228
9.2	The semantic differential technique	228
9.3	Osgood & Luria's analysis	230
9.4	Preprocessing of the data	232
9.5	Three-mode analysis	233
	<i>scale space</i>	
	<i>concept space</i>	
	<i>concept-scale interactions</i>	
9.6	Differences with Osgood & Luria	240
9.7	Concluding remarks	241
10.	ASYMMETRIC SIMILARITY DATA: ITP STUDY	243
10.1	Introduction	244
10.2	Theory, design, and data	244
	<i>theory</i>	
	<i>design and data</i>	
10.3	Theoretical subjects	246
10.4	Scale and stimulus configurations	247
10.5	Subject spaces	249
10.6	Residual/fit ratios	252
10.7	Conclusions	253
11.	SIMILARITIES AND ADJECTIVE RATINGS: COLA STUDY	255
11.1	Introduction	256
11.2	Similarity set	257
	<i>data</i>	
	<i>cola spaces</i>	
	<i>importance of dimensions</i>	
	<i>subject weights</i>	
	<i>fit of subjects</i>	
11.3	Adjective set	263
	<i>data</i>	

	<i>Schiffman et al.'s analysis</i>	
	<i>Tucker2 analysis</i>	
11.4	Similarity and adjective sets	269
11.5	Conclusion	272
12.	CORRELATION MATRICES: FOUR ABILITY-FACTOR STUDY	273
12.1	Introduction	274
12.2	Three-mode analysis of correlation matrices	
12.3	Other approaches	276
12.4	Four ability-factor study: data, hypotheses, and analyses	277
12.5	Four ability-factor study: three-mode analysis	279
	<i>test components</i>	
	<i>assessing differentiation via the core matrix</i>	
	<i>differentiation for normal children</i>	
	<i>differentiation for retarded children</i>	
12.6	Conclusion	284
13.	MULTIVARIATE LONGITUDINAL DATA: HOSPITAL STUDY	285
13.1	Introduction	286
13.2	Scope of three-mode analysis for longitudinal data	288
13.3	Analysis of data from multivariate autoregressive processes	288
	<i>introduction</i>	
	<i>component analysis of time modes</i>	
	<i>autoregressive processes</i>	
	<i>latent covariation matrix</i>	
	<i>linking autoregressive parameters to three-mode results</i>	
	<i>discussion</i>	
13.4	Growth and development of Dutch hospital organizations	298
	<i>research questions</i>	
	<i>data</i>	
	<i>results of three-mode analysis</i>	
13.5	Interpretation in terms of autoregressive processes	304
	<i>'estimation' of change phenomena and methods of analysis</i>	
	<i>checking the order of autoregressive process</i>	
	<i>assessment of change phenomena in the Hospital study</i>	
13.6	Conclusion	310
14.	GROWTH CURVES: LEARNING-TO-READ STUDY	313
14.1	Introduction	314
14.2	Data and preprocessing	315
14.3	Average learning curve	317
14.4	General characteristics of the solution	318
14.5	Analysis of interactions	319
14.6	Conclusion	323

15. THREE-MODE CORRESPONDENCE ANALYSIS: LEIDEN ELECTORATE STUDY	325
15.1 Introduction	326
15.2 Loglinear models, interactions, and chi-terms	326
15.3 Correspondence analysis for contingency tables	328
<i>two-way tables</i>	
<i>three-way tables</i>	
15.4 Leiden electorate study: data	330
15.5 Loglinear analysis	331
15.6 Model III: Political alignment of the city of Leiden	335
15.7 Model II: Decline of support for the PvdA	337
15.8 Model I: Simultaneous analysis of all non-fixed interactions	339
15.9 Conclusion	342
SAMENVATTING	345
ACKNOWLEDGEMENTS	351
APPENDICES	353
<i>classification of theoretical three-mode papers</i>	
<i>classification of applications: subject matter</i>	
<i>classification of applications: data types</i>	
<i>references to computer programs</i>	
REFERENCES	359
AUTHOR INDEX	387
SUBJECT INDEX	389

