

Meerdimensionale schaaltechnieken voor gelijkis- en keuzedata

voor Heleen en Marja

... a researcher who is looking for just *one* model that explains all responses given by a subject, may neglect the truism that there are more things between heaven and earth than dreamt of in his philosophy.

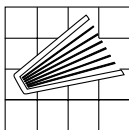
J.P. van de Geer (1995, p. 25). *Some aspects of Minkowski distances*.

— Meerdimensionale schaal- technieken voor gelijkenis- en keuzedata

**RUIMTELIJKE MODELLEN VOOR PSYCHOLOGIE,
MARKTONDERZOEK EN ANDERE WETENSCHAPPEN**

Dr. Willem A. van der Kloot

Sectie Methoden en Technieken
van Psychologisch Onderzoek
Rijksuniversiteit Leiden



UITGEVERIJ LEMMA BV - UTRECHT - 1997

Aan de totstandkoming van deze uitgave is de uiterste zorg besteed. Voor informatie die nochtans onvolledig of onjuist is opgenomen, aanvaarden auteur(s), redactie en uitgever geen aansprakelijkheid. Voor eventuele verbeteringen van de opgenomen gegevens houden zij zich gaarne aanbevolen.

Ondanks de nodige nasporingen bleek het niet mogelijk van alle opgenomen illustraties de bezitter van het copyright te achterhalen. Eventuele rechthebbenden die niet voor deze uitgave zijn benaderd, wordt verzocht zich met de uitgever in verbinding te stellen.

ISBN 90 5189 692 1
NUGI 716

© 1997 Uitgeverij LEMMA BV, Postbus 3320, 3502 GH UTRECHT

Alle rechten voorbehouden. Niets uit deze uitgave mag worden verveelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand of openbaar gemaakt in enige vorm of op enige wijze, hetzij elektronisch, hetzij mechanisch, door fotokopieën, opnamen of enige andere manier, zonder voorafgaande schriftelijke toestemming van de uitgever. Voorzover het maken van kopieën uit deze uitgave is toegestaan op grond van artikel 16b Auteurswet 1912 j° het Besluit van 20 juni 1974, Stb. 351 zoals gewijzigd bij het Besluit van 23 augustus 1985, Stb. 471 en artikel 17 Auteurswet 1912, dient men de daarvoor wettelijk verschuldigde vergoedingen te voldoen aan de Stichting Reprorrecht, Postbus 882, 1180 AW Amstelveen. Voor het overnemen van één of meer gedeelten uit deze uitgave in bloemlezingen, readers en andere compilatiewerken (artikel 16 Auteurswet 1912) dient men zich tot de uitgever te wenden.

Omslagontwerp en typografie: Twin Design BV, Culemborg

Inhoud

| | | |
|----------|---|----|
| | Voorwoord | 9 |
| 1 | Inleiding | 13 |
| | 1.1 Voorproefje: de methode van Thurstone | 13 |
| | <i>Blok 1.1 Thurstones Law of comparative judgment</i> | 18 |
| | 1.2 Kenmerken van schaaltechnieken | 27 |
| | 1.3 Opzet van dit boek | 33 |
| 2 | MDS als cartografisch probleem | 35 |
| | 2.1 De kaart van Nederland | 35 |
| | 2.2 De Young-Householdermethode | 36 |
| | <i>Blok 2.1 Scalaire producten met de centroïde als oorsprong</i> | 46 |
| | <i>Blok 2.2 Het probleem van de additieve constante</i> | 56 |
| 3 | Datatheorie | 63 |
| | 3.1 Nabijheid en afstand | 63 |
| | 3.2 De datatheorie van Coombs | 64 |
| | 3.3 Het systeem van Carroll, Arabie en Young | 67 |
| | 3.4 Kruskals typologie van data | 73 |
| | 3.5 Een indeling van schaalproblemen | 75 |
| | 3.6 Transformaties | 79 |
| | <i>Blok 3.1 Het vinden van monotone transformaties</i> | 83 |
| | 3.7 Recapitulatie | 86 |

| | | |
|----------|--|-----|
| 4 | Afstandsfuncties en afstandsmodellen | 89 |
| 4.1 | Afstandsfuncties | 89 |
| 4.2 | Profielafstanden | 96 |
| 4.3 | Afstandsfuncties als model | 99 |
| | <i>Blok 4.1 Afstandsmodellen als onderwerp van onderzoek</i> | 101 |
| | <i>Blok 4.2 Psychologische interpretaties van Minkowski's p</i> | 105 |
| 4.4 | Recapitulatie | 106 |
| 5 | Varianten van ruimtelijke afbeeldingen | 109 |
| 5.1 | Transformaties van ruimtelijke afbeeldingen | 109 |
| 5.2 | Afstandsbehoudende transformaties | 110 |
| 5.3 | Transformaties die niet afstandsbehoudend zijn | 116 |
| 5.4 | Implicaties | 118 |
| 6 | Niet-metrische MDS: iteratieve algoritmen | 121 |
| 6.1 | De Shepard-Kruskalmethode | 121 |
| 6.2 | Iteratieve algoritmen voor MDS | 122 |
| | <i>Blok 6.1 Het verbeteren van de configuratie</i> | 131 |
| 6.3 | Andere algoritmen | 139 |
| 6.4 | ALSCAL | 141 |
| | <i>Blok 6.2 Niet-metrische analyse van de afstanden tussen acht steden</i> | 148 |
| 7 | Interpretatie van MDS-oplossingen | 165 |
| 7.1 | Aspecten van oplossingen | 165 |
| 7.2 | Dimensionaliteit en stress | 166 |
| | <i>Blok 7.1 Monte Carlo-methoden voor het bepalen van dimensionaliteit</i> | 172 |
| 7.3 | De inhoudelijke interpretatie van MDS-oplossingen | 178 |
| 7.4 | Externe analyse | 183 |
| 7.5 | Procrustesanalyse | 196 |
| 8 | Asymmetrische nabijheidsgegevens | 199 |
| 8.1 | Uitbreidingen van CMDS | 199 |
| 8.2 | Asymmetrische gelijkennisgegevens | 200 |
| 8.3 | Modellen met parameters voor asymmetrie | 203 |
| | <i>Blok 8.1 Nederland op de fiets: wind mee en wind tegen</i> | 211 |

| | | |
|---------------|--|---------|
| 9 | MDS van drieweg/tweemodale data | 223 |
| 9.1 | MDS van vierkante driewegdata | 223 |
| | <i>Blok 9.1 Het aggregeren van ordinale nabijheidsgegevens</i> | 224 |
| 9.2 | MDS met het INDSCAL-model | 228 |
| 9.3 | Interpretatie van een INDSCAL-oplossing | 231 |
| | <i>Blok 9.2 RMDS en INDSCAL: veronderstelde relaties tussen landen</i> | 238 |
| 9.4 | Andere individuele-verschillenmodellen | 249 |
| 9.5 | Asymmetrische drieweg/tweemodale nabijheidsdata | 250 |
| 10 | Keuzedata | 251 |
| 10.1 | Rechthoekige nabijheidsmatrices | 251 |
| 10.2 | MDS van keuzedata | 253 |
| 10.3 | Keuzedata als nabijheden | 254 |
| 10.4 | HOMALS | 257 |
| | <i>Blok 10.1 Een Homals-analyse van keuzedata</i> | 264 |
| 10.5 | Keuzedata als dominantierelaties | 272 |
| 11 | Preferentierangordeningen | 285 |
| 11.1 | Preferentierangordeningen | 285 |
| 11.2 | Metrische analyse met het vectormodel | 286 |
| | <i>Blok 11.1 Metrische analyse van preferentiedata</i> | 289 |
| 11.3 | Niet-metrische analyse met het vectormodel | 295 |
| | <i>Blok 11.2 Een voorbeeld van PRINCALS</i> | 296 |
| | <i>Blok 11.3 PRINCALS op de rangordeningen van auteursbijdragen</i> | 310 |
| 11.4 | Correspondentieanalyse | 311 |
| | <i>Blok 11.4 Misdaad en straf</i> | 313 |
| | <i>Blok 11.5 Een inhoudsanalyse van kookboeken</i> | 317 |
| 11.5 | Het ontvouwingsmodel | 319 |
| | <i>Blok 11.6 Eendimensionaal ontvouwen</i> | 324 |
| 11.6 | Numerieke oplossingen voor het ontvouwingsprobleem | 328 |
| | <i>Blok 11.7 Ontvouwen met HOMALS: paarsgewijze coderingen</i> | 336 |
| | <i>Blok 11.8 Unfolding met HOMALS: de methode van Heiser</i> | 340 |
| | <i>Blok 11.9 HOMALS op tot keuzedata herleide rangordeningen</i> | 343 |

| | | |
|-----------|---|-----|
| 12 | De opzet van een MDS-onderzoek | 347 |
| 12.1 | Aspecten van een MDS-onderzoek | 347 |
| 12.2 | Vraagstelling | 348 |
| 12.3 | Stimuli en condities | 349 |
| | <i>Blok 12.1 Facet-theorie</i> | 352 |
| 12.4 | Selectie van respondenten | 355 |
| 12.5 | Dataverzamelingsprocedures | 356 |
| 12.6 | Beoordelingen door proefpersonen | 359 |
| | <i>Blok 12.2 Consistentie en circulaire triaden</i> | 364 |
| | <i>Blok 12.3 Kelly's repertory grid</i> | 367 |
| 12.7 | Tweeweg/tweemodale gegevens | 371 |
| 12.8 | De uitvoering van de analyses | 375 |
| 12.9 | Interpretatie | 379 |
| | <i>Blok 12.4 De externe interpretatie van waarden</i> | 379 |
| 12.10 | Rapportage | 380 |
| 12.11 | Simulatie | 381 |
| 13 | Bijzondere toepassingen | 387 |
| 13.1 | Bijzondere observaties | 387 |
| 13.2 | MDS van paarsgewijze preferentiedata | 387 |
| 13.3 | De analyse van sorteergegevens | 392 |
| | <i>Blok 13.1 SPSS PROXIMITY</i> | 397 |
| | <i>Blok 13.2 Correspondentieanalyse van G</i> | 403 |
| 13.4 | MDS-analyse van correlatiecoëfficiënten | 411 |
| | <i>Blok 13.3 Principale-componentenanalyse en MDS</i> | 415 |
| 13.5 | Drieweg/driemodale data | 418 |
| | <i>Blok 13.4 INDSCAL op profielafstanden</i> | 425 |
| 14 | Slotopmerkingen | 427 |
| | Appendix: Basiskennis matrixalgebra | 431 |
| | Literatuur | 439 |
| | Register | 450 |

— Voorwoord

Dit boek komt voort uit een reeks cursussen *Meerdimensionale schaaltechnieken* (MDS) die ik vanaf 1978 binnen de studierichting Psychologie aan de Rijksuniversiteit Leiden, meestal alleen, soms met anderen, verzorgd heb. In die cursussen heb ik tot een jaar of vier geleden gebruik gemaakt van verschillende leerboeken. Achtereenvolgens waren dat Kruskal en Wish (1978), Schiffman, Reynolds en Young (1981), Coxon (1982) en Young en Hamer (1987) en van al die leerboeken zijn er wel enkele kenmerken terug te vinden in dit boek. Een handig gebruik dat ik van Young (Schiffman e.a., 1981; Young & Hamer, 1987) heb overgenomen is de notatie om MDS-gegevens met de O van observaties aan te duiden, in plaats van met de meer gebruikelijke symbolen Δ en δ . Ook tot Young te herleiden is de tamelijk uitgebreide behandeling van de eigenschappen van ruimtelijke afbeeldingen (Hoofdstuk 5) en de nadruk op de begrippen transformaties en conditionaliteit (Hoofdstuk 3, Hoofdstuk 6). Het leerboek van Coxon heeft vooral als voorbeeld gediend bij het streven naar een zekere volledigheid en praktische bruikbaarheid.

Behalve dat dit werk beïnvloed is door andere boeken, zijn er ook invloeden geweest van collega's met wie ik in de loop der tijd samen cursussen over MDS gegeven heb. Mijn allereerste MDS-cursus was een gezamenlijke onderneming met prof.dr. Willem Heiser. Ook die eerste cursus begon met de Thurstone-methode; voorzover ik mij herinner, misschien wel met hetzelfde misdaandonderzoek dat in Hoofdstuk 1 besproken wordt. De richting waarin latere cursussen en uiteindelijk dit boek zich ontwikkeld hebben, is voor een deel tot die eerste cursus te herleiden. Ook anderen - ex studentassistenten, nu collega's - hebben mede vormgegeven aan 'mijn' MDS-cursussen. Met name de manier waarop Jacques Commandeur het vectormodel uitlegde, heeft als voorbeeld gediend voor wat ik daarover in dit boek (Hoofdstuk 11) geschreven heb.

Mijn eerste kennismaking met MDS vond plaats in de cursus die prof.dr. John van de Geer in 1966 over dit onderwerp gaf. Bij zijn uitleg van Coombs' data-theorie gebruikte hij als voorbeeld erwten en worteltjes. De door mij genoemde spruitjes en andijvie zijn dan ook een *hommage* aan Van de Geer. Zijn cursus liet zien dat ook het dagelijks leven uit relaties tussen punten bestaat. Ten slotte zijn er in dit boek invloeden terug te vinden van het werk van het Meerling-collectief (1988, 1989). Enkele stukjes die ik destijds voor dit gezelschap geschreven en geredigeerd heb, zijn in dit boek geparafraseerd overgenomen.

Wat zijn nu de belangrijkste kenmerken van dit boek? De eerste zeven hoofdstukken bevatten een uiteenzetting van de verschillende soorten data die met MDS geanalyseerd kunnen worden, de verschillende definities van het begrip afstand, de verschillende meetniveaus en toegestane transformaties, het basis-algoritme van metrische en niet-metrische MDS, afstandsbehoudende en andere transformaties van ruimtelijke configuraties, en de interpretatie van MDS-configuraties en het INDSCAL-model. Deze onderwerpen in de een of andere vorm vaak wel terug te vinden in de meeste andere leerboeken over MDS, zij het dat daarin niet al deze onderwerpen behandeld worden of niet alle onderwerpen even uitgebreid aan de orde komen. Het eigene van dit boek komt vooral naar voren in Hoofdstuk 8 over MDS op asymmetrische data en in de Hoofdstukken 10 tot en met 13 waarin de analyse van keuzedata en preferentierangordeningen behandeld worden. Voor de analyse van dit soort data worden de niet-lineaire multivariate technieken HOMALS, PRINCALS en ANACOR (zie Gifi, 1990) aanbevolen. Ook wordt in dit kader de Mokschaal-analyse besproken. Ten slotte wordt in Hoofdstuk 13 ook drieweg principale-componentenanalyse kort genoemd, na een uitgebreide uiteenzetting over de analyse van sorteergegevens en een wat kortere verhandeling over MDS van correlatiecoëfficiënten.

Waar ik bij het schrijven veel plezier aan beleefde, was het gebruik van voorbeelden, data en resultaten van 'echt' onderzoek dat min of meer 'in mijn omgeving' had plaatsgevonden of dat ik ooit zelf, al dan niet samen met (oud)studenten en collega's, verricht heb. Een van die studies - over auteurschap en auteursvolgorde (Van der Kloot & Willemsen, 1991) - loopt als een *running gag* door dit hele boek heen: de betreffende data worden op verschillende manieren geanalyseerd, wat meestal wel, maar - en dat is leerzaam! - niet altijd dezelfde resultaten oplevert. De tweede *running gag* is de kaart van Nederland die op verschillende plaatsen opnieuw wordt 'uitgerekend'.

Dat ik dit boek alléén heb geschreven, neemt niet weg dat er een aantal personen is die op enigerlei wijze aan vorm en inhoud hebben bijgedragen. Dezen wil ik daarvoor dan ook bedanken. In de eerste plaats zijn dat mijn collega's met wie ik ooit het onderzoek verricht heb dat in dit boek als voorbeeldmateriaal gebruikt wordt. Dank ben ik eveneens verschuldigd aan Femke Heide die de externe analyses uitvoerde die in Hoofdstuk 7 besproken worden. Ook wil

ik hier de studenten bedanken die in 1992, 1994 en 1995 de vele tik- en andere fouten uit de eerste manuscriptversies van dit boek haalden en de tekst hier en daar van commentaar voorzagen. In het bijzonder dank ik hier Adriënne Rietbergen, Mark Steverink, Hans Brinkman en Gerard Hollemans voor hun constructieve bijdragen. Vooral Inge van der Valks wekelijkse lijstjes met 'foutjes' betekenden een enorme hulp.

Schrijven wordt - zoals veel andere activiteiten in het leven - soms vergeleken met het maken van een reis: je vertrekt uit een bekend beginpunt om via lange omzwervingen op een onvoorziene bestemming aan te komen. Voor dit boek is die voorstelling veel te romantisch. Voorzover het schrijven van een boek als dit op reizen lijkt, lijkt het op een vliegreis. Na een aantal hoofdstukken komt er een *point of no return*; stoppen is dan onmogelijk. Dat ik dit point of no return bereikte, is mede te danken aan het stimulerende interim management van Eric van der Loo. En niet in het minst aan Tineke, Marja en Heleen.

Willem A. van der Kloot

