

SAMENVATTING

Met dit boek worden drie doelen nagestreefd. In de eerste plaats is het een monografie over drieweg-hoofdassenanalyse, en er is gepoogd alle belangrijke aspecten van deze techniek te bespreken. Verreweg het grootste gedeelte van de literatuur op dit terrein, alsmede aanverwante gebieden uit de meerdimensionale schaalmethoden worden (kritisch) besproken.

In de tweede plaats worden verbeterde methoden gepresenteerd om drieweganalyse uit te voeren, alsmede de consequenties hiervan doorgelicht. Ook worden theoretische bijdragen gepresenteerd over het transformeren van kernmatrices naar wat 'simpele structuren' worden genoemd.

Het verschaffen van een handleiding van de methoden in de praktijk van het sociaal-wetenschappelijk onderzoek kan worden aangemerkt als een derde doel bij het schrijven van dit boek. Zowel in- als uitvoerproblemen en interpretaties worden in het algemeen besproken en in detail toegelicht bij de analyse van een differentiaal voorbeelden. Om de bruikbaarheid voor de praktijk te verhogen is een (vrijwel compleet) overzicht gemaakt van toepassingen van drieweganalyse, die bovendien geclassificeerd zijn naar inhoudelijke onderwerpen, zodat een ieder aansluiting kan zoeken bij vakgenoten die eerder deze techniek hebben toegepast.

De eerste twee hoofdstukken hebben een algemeen karakter. Hoofdstuk 1 is een organisatorische wegwijzer voor het boek. Het bevat een leeswijzer, een woordenlijst en een overzicht van de gebruikte notatie. Hoofdstuk 2 is het boek zelf in een notedop en kan worden opgevat als een geannoteerde inhoudsopgave of als een

handreiking voor diegenen die geen zin of tijd hebben om het hele boek door te lezen.

De hoofdstukken uit Deel I richten zich op de theoretische aspecten verbonden met drieweg-hoofdassenanalyse. In Hoofdstuk 3 worden de modellen (het Tucker3 en Tucker2 model) die de basis vormen voor de rest van het boek in de context geplaatst van andere modellen die op dit terrein zijn voorgesteld. Deze laatste modellen kunnen in twee klassen worden onderverdeeld, namelijk hoofdassenmodellen en factoranalyse-modellen, waarbij de laatste één stochastische weg hebben en bij de eerste klasse alle wegen niet-stochastisch zijn. De klasse van hoofdassenmodellen kan weer worden onderverdeeld in twee subklassen, namelijk die met drie 'ingedikte' wegen, zoals het Tucker3 model, drieweg-schaling, PARAFAC1 en INDSCAL, en die met twee ingedikte wegen zoals het Tucker2 model, PARAFAC2, IDIOSCAL, CANDECOMP en INDSCAL. Dit soort modellen wordt doorgaans geanalyseerd met behulp van alternerende kleinste-kwadraatmethoden, terwijl de factoranalyse-modellen worden geanalyseerd met technieken uit het domein van covariantie-structuurmodellen. Hoofdstuk 3 besluit met de bespreking van een aantal mogelijke toevoegingen aan de Tucker modellen, zoals schattingsprocedures over ontbrekende gegevens, optimale schaalprocedures voor gegevens van lagere meetniveaus, faciliteiten om externe analyses uit te voeren, uitbreidingen naar n wegen, etc.

Hoofdstuk 4 bevat de (technische) kern van het boek. Alternerende kleinste-kwadraatmethoden (ALS) voor de Tucker modellen worden gepresenteerd. Daarbij komen aan de orde: het bestaan van exacte en benaderende oplossingen, constructie van ALS-algoritmen, hun convergentie, alsmede een aantal andere technische details. Drie bestandjes worden gebruikt om de nauwkeurigheid en de correctheid van de algoritmen te bestuderen en om na te gaan hoe gevoelig de oplossingen zijn voor toenemende toevalsfluctuaties in de gegevens.

Deel I eindigt met een bespreking in Hoofdstuk 5 van methoden om kernmatrices zo te transformeren dat ze een eenvoudige structuur krijgen, bijvoorbeeld met veel nullen. Twee algoritmen om dit te bereiken worden besproken; de ene gebaseerd op orthonormale trans-

formaties en de andere op niet-singuliere transformaties. De problemen wat betreft de laatste methode worden gesignaleerd. Twee bestanden worden gebruikt om de methoden te illustreren en te evalueren.

Ook Deel II bevat voornamelijk theoretische verhandelingen, maar nu ligt de nadruk op die problemen die voortkomen uit het toepassen van drieweg-hoofdassenanalyse in de praktijk. Drie gebieden worden benadrukt: het voorbereiden van de oorspronkelijke gegevens zodat ze geschikt zijn voor een drieweganalyse, het bewerken van de ruwe uitvoer zodat die gemakkelijk te interpreteren valt, en het analyseren van het gedeelte van de gegevens wat niet overeenkomt met het gebruikte drieweg-model.

Het eerste deel van Hoofdstuk 6 geeft een overzicht van voorstellen die gedaan zijn om ruwe gegevens zo te bewerken dat ze geschikt zijn voor een (drieweg-)hoofdassenanalyse. Met name methoden om verschillen in gemiddelden en varianties te elimineren worden behandeld; hierbij wordt een verschil gemaakt tussen (on)interpreteerbare en (on)vergelijkbare gemiddelden en varianties. Een aantal modellen die gebruik maken van interpreteerbare gemiddelden en varianties worden besproken, met name die welke bestaan uit additieve termen voor de gemiddelden, multiplicatieve termen voor de varianties en product-termen voor de componenten. Ook wordt het probleem van iteratieve standaardizatie aangestipt.

In het tweede deel van Hoofdstuk 6 wordt er aandacht geschonken aan het interpreteren van de resultaten van een drieweganalyse en aan methoden om deze interpretatie te vergemakkelijken en te verbeteren. Hierbij passeren het schalen en interpreteren van componenten en de kerumatrix de revue, alsmede het gebruik van 'gezamenlijke componentenruimten' en hun afbeeldingen, en verder het gebruik van component scores in drieweganalyse.

In Hoofdstuk 7 wordt de functie en het nut van residuenanalyse besproken voor (drieweg-)hoofdassenanalyse. Ook worden procedures en gedetailleerde aanbevelingen voor het behandelen van drieweg-residuen gepresenteerd en geïllustreerd. Met name komen aan bod: variantieanalyse op gekwadrateerde residuen, kwadratensommengrafieken, en het gebruik van normale kansverdelingsgrafieken.

In Deel II wordt de theorie uit de voorgaande delen toegepast en toegelicht met behulp van diverse bestanden. Elk van de geanalyseerde bestanden is opgenomen als een vertegenwoordiger van een bepaalde klasse toepassingen; onderzoekers met soortgelijke gegevens kunnen mogelijk hieruit inspiratie putten voor een geschikte aanpak voor hun eigen analyse.

Hoofdstuk 8 (*Attachment study*) bevat een gedetailleerde analyse van de reactie van kleine kinderen op een gestandaardiseerde procedure die de aard en de mate van gehechtheid van die kinderen aan hun moeder poogt te meten. De gegevens, scores van individuen op een aantal variabelen onder verschillende condities, zijn kenmerkend voor het soort gegevens dat met vrucht via drieweganalyse kan worden onderzocht.

Hoofdstuk 9 (*Triple personality study*) bevat een voorbeeld van gegevens verzameld met behulp van semantische differentialen, waarbij met name de individuele verschillen van groot belang zijn. Verschillen in gebruik van de relaties tussen schalen en begrippen staan centraal in deze analyse. De gegevens zijn afkomstig van één vrouw met drie verschillende persoonlijkheden: Eve Black, Eve White en Jane. De behandeling van de gegevens kan als voorbeeld dienen voor andere onderzoeken waarbij vergelijkbare beoordelingsschalen of testen worden gebruikt.

Hoofdstuk 10 (*ITP study*) bevat een voorbeeld van wat asymmetrische gelijkenisgegevens zou kunnen worden genoemd. In plaats van een gedetailleerde behandeling van de gegevens wordt dit hoofdstuk gebruikt om de 'theoretische proefpersoon' te introduceren, dat wil zeggen een proefpersoon geconstrueerd op basis van theoretische inzichten. Zo'n persoon is met name nuttig om als leidraad te gebruiken bij de interpretatie van de hoofdassen in een proefpersoonsruimte.

Hoofdstuk 11 (*Cola study*) bevat een heranalyse van gegevens over de smaak van cola's, waarbij met name aandacht wordt besteed aan de vergelijking van gelijkenisoordelen en oordelen verzameld met beoordelingsschalen. De resultaten van de drieweganalyse op de gelijkenisoordelen worden vergeleken met die van een drietal schaalmethoden voor individuele verschillen, aangezien dit soort gegevens doorgaans met deze methoden wordt geanalyseerd.

Hoofdstuk 12 (*Four ability-factor study*) laat zien hoe drieweganalyse kan worden gebruikt om correlatiematrices te analyseren. Hiermee wordt een mogelijkheid aangegeven om transversale gegevens eventueel uit verschillende bronnen te onderzoeken. Vooral in die gevallen waarin slechts de correlatiematrices gepubliceerd zijn, maar niet de oorspronkelijke gegevens kan een dergelijke analyse nuttig zijn. Het voorbeeld behandelt de analyse van matrices met correlaties tussen verschillende intelligentietesten, die afkomstig zijn van normale en zwakzinnige kinderen van verschillende leeftijden.

Hoofdstuk 13 (*Hospital study*) behandelt de analyse van multivariate longitudinale gegevens met behulp van drieweganalyse, geïllustreerd met een voorbeeld uit de organisatiesociologie. In de studie worden de structuren in de organisatie van Nederlandse ziekenhuizen bekeken over de jaren heen. Verschillende problemen met betrekking tot de analyse van dergelijke gegevens worden besproken, met name de relatie tussen drieweganalyse en autoregressie-modellen.

Hoofdstuk 14 (*Learning-to-read study*) bevat een uitbreiding van Tucker's werk over leercurven. De leercurven in deze studie zijn afkomstig van kinderen die in de loop van een jaar leren lezen en ze worden gevormd door de scores op een aantal testen die hun leesvordering meten. Een vluchtige vergelijking met resultaten van een lineaire logistische analyse op dezelfde gegevens wordt gemaakt.

Hoofdstuk 15 (*Leiden electorate study*) bevat een wat afwijkend voorbeeld, aangezien het gaat om de analyse van residuen uit kruistabellen met frequenties, in plaats van een analyse van gemeten gegevens. De uitslagen van Leidse kiesdistricten in drie verschillende verkiezingen vormen de basis voor de analyse. Allereerst worden de relevante interacties in de gegevens bepaald met loglineaire analyse; vervolgens zijn de residuen van een aantal loglineaire modellen onderzocht met behulp van drieweganalyse. De gebruikte procedure is een drieweg-generalisatie van correspondentieanalyse.

