

MULTIVARIATE ANALYSE MET BEHULP VAN PC-MDS 5.1

Tammo Bijmolt, Edward Rosbergen en Marco Vriens*

Samenvatting

In dit artikel wordt een beschrijving en een beoordeling gegeven van PC-MDS versie 5.1. Dit is een pakket bestaande uit 22 modules voor multidimensional scaling en een aantal overige multivariate analyses. Na een korte inleiding worden de modules afzonderlijk besproken. De toepassing van het pakket wordt verduidelijkt met een voorbeeld. Daarna wordt ingegaan op de handleiding en een boek over het pakket. Tevens wordt deze nieuwe versie vergeleken met versie 4.1 van PC-MDS. Ten slotte volgt een evaluatie van PC-MDS.

* De auteurs zijn verbonden aan de R.U. Groningen, Faculteit der Economische Wetenschappen, Vakgroep Bedrijfseconomie, Sectie Marktkunde & Marktonderzoek. Adres: Postbus 800, 9700 AV, Groningen, Tel.: 050-636328.

1. Administratieve en technische gegevens

Naam	PC-MDS versie 5.1
Leverancier	Scott M. Smith, Ph.D. Department of Marketing Institute of Business Management Brigham Young University Provo, Utah 84602, Verenigde Staten
Prijs	\$ 500 voor nieuwe gebruikers en \$ 100 voor een update vanaf versie 4.1 en \$ 200 voor een update vanaf een oudere versie
Omvang programma	22 modules samen met voorbeeld databestanden op 10 floppy's van elk 360 KB
Hardware	IBM (of compatibel) PC, XT of AT, harde schijf niet noodzakelijk
Geheugenruimte	ten minste 512K
Overige	Geen mathematische coprocessor of grafische kaart vereist. Een kleurenscherm verbetert de schermweergave van PC-MDS aanzienlijk.

2. Inleiding

PC-MDS is een verzameling van reeds bestaande computerprogramma's, die geschreven zijn in Microsoft Fortran 77. In 1990 is een nieuwe versie (5.1) van dit pakket verschenen. De naam PC-MDS doet vermoeden dat het hier om typische scaling programma's gaat. Echter, de afkorting MDS staat voor MultiDimensional Statistics. Naast scaling programma's, zoals KYST en INDSCAL, zijn er modules opgenomen voor onder andere conjuncte analyse, clusteranalyse, correspondentieanalyse en factoranalyse. Tabel 1 geeft een overzicht van de modules die opgenomen zijn in PC-MDS, met daarbij vermeld in welke sectie de betreffende modules zullen worden besproken. Voor de meeste modules hoeft alleen een databestand te worden aangemaakt. De parameterwaarden moeten dan in de eerste regels van het databestand worden opgenomen, of via menu's in het programma worden aangegeven. In de nieuwste versie van PC-MDS moet voor vier modules, te weten DISCRIM, FACTOR, REGRESS en FREQ, tevens een opdrachtenbestand worden gemaakt.

In het navolgende zullen we de mogelijkheden van PC-MDS kort de revue laten passeren en op beknopte wijze aangeven waar de sterke en zwakke plekken van het pakket liggen. We bespreken in sectie 3 de verschillende modules van PC-MDS, waarbij we iets langer stil zullen staan bij de ons inziens belangrijkere modules. In sectie 4 bespreken we een voorbeeld van een analyse uitgevoerd met behulp van KYST, PROFIT, CORRESP en FMATCH. Daarna, in sectie 5, gaan we in op de handleiding. Bij het boek van Green, Carmone en Smith (1989) wordt een deel van het PC-MDS pakket geleverd. In sectie 6 gaan we kort in op deze versie van PC-MDS en het betreffende boek. Vervolgens worden versie 4.1 en 5.1 met elkaar vergeleken in sectie 7. Tenslotte volgt in sectie 8 een algemene evaluatie van het pakket PC-MDS.

multidimensional scaling (3.1)		conjuncte analyse (3.2)	
KYST	multidimensional scaling	MONANOVA	monotone ANOVA
INDSCAL	analyse van individuele verschillen	TRADEOFF	trade-off conjuncte analyse
PREFMAP	externe preferentie-analyse	CONJOINT	conjuncte analyse met OLS
MDPREF	interne preferentie-analyse		
PROFIT	property fitting		
clusteranalyse (3.3)		correspondentieanalyse (3.4)	
CLUSTER	Howard-Harris cluster-analyse	CORAN	Lebart, Morineau en Warwick algoritme
HICLUST	hiërarchische cluster-analyse	CORRESP	Carroll, Green en Schaffer algoritme
WARD	Ward's agglomeratieve clusteranalyse		
multivariate analyse (3.5)		overige modules (3.6)	
DISCRIM	discriminantanalyse	FMATCH	procrustesanalyse
FACTOR	factoranalyse	DISTRAN	afstanden berekenen
AID	automatische interactie detectie	CASE5	Thurstone's Case V
REGRESS	regressieanalyse	FREQ	frequentietabellen
		EDIT	ASCII editor

Tabel 1. De modules in PC-MDS.

3. Mogelijkheden van PC-MDS

In deze sectie worden de modules van PC-MDS stuk voor stuk besproken. Daarbij geldt dat de indeling van de modules in deze sectie overeenkomt met de in tabel 1 weergegeven indeling.

3.1. Modules voor multidimensional scaling

Een belangrijk gedeelte van PC-MDS is gewijd aan multidimensional scaling. Met behulp van deze technieken kunnen gelijkennisgegevens en preferentiegegevens worden afgebeeld in een

meerdimensionale ruimte. In PC-MDS zijn 5 modules opgenomen voor het uitvoeren van een multidimensional scaling analyse.

KYST

KYST staat voor "Kruskal, Young, Shepard en Torgerson" en is een samenvoeging van MDSCAL (Kruskal, 1964 a en b) en TORSCA (Young en Torgerson, 1967). KYST is in staat om preferenties of ongelijkheden te vertalen naar een figuur waarin de stimuli staan afgebeeld. Binnen KYST zijn legio mogelijkheden aanwezig wat betreft de vorm van de in te voeren data-matrix en de uit te voeren bewerkingen op deze data-matrix. Naast deze vele mogelijkheden kent KYST echter ook zijn beperkingen. Zo kunnen er maximaal 100 objecten worden geschaald op ten hoogste 10 dimensies. Het totale aantal gegevens over deze objecten mag de 5000 niet overschrijden. Ook is het niet mogelijk het programma meerdere malen achter elkaar te laten lopen en de achtereenvolgende uitvoeren naar verschillende files te schrijven. Daarnaast kan men bij het veranderen van de parameters van het programma voor verrassingen komen te staan. Zo zal men er rekening mee moeten houden dat de parameteropties in de handleiding niet overeen hoeven te komen met die in het programma. Dit geldt bijvoorbeeld voor de te kiezen afstandsfunctie. Volgens de handleiding kan hier een waarde tussen 0 en oneindig worden ingevoerd, terwijl het programma alleen de waarden 1 en 2 accepteert.

INDSCAL

INDSCAL is een programma dat ontwikkeld is om de individuele verschillen tussen meerdere subjecten te analyseren (Carroll en Chang, 1970). Een verschil tussen INDSCAL en KYST is dat bij INDSCAL de gegevens op individueel niveau worden geanalyseerd, terwijl dit bij KYST op geaggregeerd niveau gebeurt. Bovendien is INDSCAL veel minder flexibel ten aanzien van de invoer. Een ander verschil met KYST is het feit dat INDSCAL niet beschikt over menu-schermen. Daarom moet de eerste regel van de invoer-file gewijzigd zijn aan de in te stellen opties. Binnen INDSCAL kunnen maximaal 100 data-sets van individuen worden ingelezen, waarbij per data-set maximaal 45 stimuli met elkaar vergeleken kunnen worden.

PREFMAP

Deze module stelt de gebruiker in staat een stimulusruimte te relateren aan preferentiedata (externe ontvouwing). Hierbij kan de stimulusruimte zijn verkregen met behulp van KYST of INDSCAL. Vervolgens voegt PREFMAP hier de individuen op basis van hun preferenties aan toe. De analyse vindt plaats op basis van vier, hiërarchisch geordende modellen. Het maximale aantal individuen dat bij de analyse betrokken kan worden, bedraagt 40. Deze individuen worden in een ruimte geplaatst die maximaal 5-dimensionaal mag zijn en waarin zich ten hoogste 150 stimuli (volgens de handleiding, volgens het programma 200) mogen bevinden. Bij elke multidimensional scaling-module is het mogelijk een criterium voor nauwkeurigheid van de oplossing in te stellen. Bij alle andere modules voor multidimensional scaling van PC-MDS moet de gewenste waarde worden ingevuld, alleen bij PREFMAP moet dat door middel van een code worden gedaan (bijv. $1=0.001$). Hier wordt in de handleiding echter geen melding van gemaakt.

MDPREF

MDPREF staat voor MultiDimensional Preference Scaling. MDPREF is een metrisch model dat gebaseerd is op principale componentenanalyse. Zoals de titel al suggereert, kunnen met

MDPREF voorkeuren van subjecten met betrekking tot de diverse stimuli worden geanalyseerd (interne ontvouwing). Het doel van MDPREF is een figuur te maken waarin de subjecten als vectoren en de stimuli als punten staan afgebeeld. Met MDPREF kunnen maximaal 100 stimuli, beoordeeld door ten hoogste 100 subjecten, worden geanalyseerd.

PROFIT

Het doel van PROFIT is de met behulp van bijvoorbeeld KYST of INDSICAL gevonden stimulusruimte te relateren aan attributen/eigenschappen. De eigenschappen worden door middel van vectoren in de stimulusruimte afgebeeld door PROFIT. Voor de bepaling van de ligging van de vectoren kan PROFIT gebruik maken van twee procedures, te weten: een procedure die gebaseerd is op lineaire regressie, of een procedure die gebaseerd is op het optimaliseren van een index van niet-lineaire correlatie. Met behulp van PROFIT kunnen maximaal 60 eigenschappen in de stimulus-ruimte worden afgebeeld. Deze ruimte mag ten hoogste 10-dimensionaal zijn en er mogen zich niet meer dan 400 stimuli in bevinden.

3.2. Modules voor conjuncte analyse

De drie modules van PC-MDS waarmee een conjuncte analyse kan worden uitgevoerd verschillen sterk van elkaar. Deze verschillen betreffen zowel het model als de mogelijk invoer.

MONANOVA

Deze module bevat het MONANOVA-programma van Kruskal en Carmone (1969). Met behulp van dit programma kan men conjuncte analyses uitvoeren. Het programma kent echter een aantal beperkingen. In de eerste plaats kunnen alleen additieve hoofdeffecten modellen geschat worden. In de tweede plaats kunnen alleen data geanalyseerd worden die verzameld zijn met behulp van full factorial designs. Technisch is het programma beperkt tot 12 attributen, met maximaal 100 niveaus per attribuut. Het totale aantal datapunten is beperkt tot 4000. In de praktijk van conjuncte analyse studies zijn deze beperkingen slechts theoretisch van aard omdat in dergelijke studies meestal niet meer dan 5 attributen meegenomen worden in verband met het probleem van information overload (Vriens en Wittink, 1990). Ook het gemiddeld aantal niveaus per attribuut is in de praktijk beperkt tot ongeveer 5 omdat anders het aantal door de respondent te evalueren al snel te groot (Vriens en Wilms, 1991). De echte beperking van MONANOVA voor gebruik in de praktijk is het feit dat alleen full factorial designs geanalyseerd kunnen worden.

TRADEOFF

De TRADEOFF-module is een programma voor nonmetrische conjuncte analyse, dat ontwikkeld is door Nehls, Seaman en Montgomery (1974). Met dit programma is het mogelijk om data die verzameld zijn met behulp van trade-off matrices (zie o.a. Vriens en Wittink, 1990) te analyseren. In de praktijk van het marktonderzoek wordt deze methode nauwelijks meer toegepast (Wittink en Cattin, 1989). Dit is overigens niet geheel terecht, omdat met name in situaties met meer dan 5 attributen aangeraden wordt om de trade-off dataverzamelmethode te gebruiken in plaats van de full-profile methode (Green en Srinivasan, 1978). Met behulp van de TRADEOFF-module kunnen zowel additieve als multiplicatieve modellen geschat worden. Technisch is de module beperkt tot maximaal 200 attributen, met maximaal 10 niveaus.

CONJOINT

De CONJOINT-module is gebaseerd op OLS-regressie. Dit betekent dat in de conjuncte analyse de afhankelijke variabele opgevat wordt als een variabele die op intervalniveau gemeten is. Uitgaande van full-profile data, kunnen alleen additieve modellen geschat worden. De factorial designs worden omgezet in designmatrices om vervolgens de OLS-schatter te berekenen. Daarnaast kent CONJOINT een simulator die preferentieaandelen schat gebruikmakend van het "eerste-keuze" model en het "gemiddelde-keuze" model. Het maximum aantal profielen is 64, terwijl het aantal factoren is beperkt tot 40 met een totaal maximaal aantal niveaus van 80.

3.3. Modules voor clusteranalyse

Er zijn binnen PC-MDS drie modules beschikbaar om een clusteranalyse uit te voeren. Dit zijn CLUSTER, HICLUST en WARD. De modules HICLUST en WARD bevatten hiërarchische algoritmen, CLUSTER bevat daarentegen een nonhiërarchisch optimaliseringsalgoritme. De optimaliseringsmethoden zijn over het algemeen beter in het terugvinden van vooraf bekende clusters dan de hiërarchische methoden. Bovendien presteert het Ward algoritme beter dan de overige hiërarchische clusteralgoritmen (Punj en Stewart, 1983).

CLUSTER

Met deze module kan een Howard-Harris clustering worden uitgevoerd. Gegeven het aantal clusters worden met behulp van de 'K-means methode' de subjecten (maximaal 3000) over de cluster verdeeld zodanig dat de variantie binnen de clusters minimaal is. De uiteindelijke clusteroplossing is in ieder geval een lokaal optimum. Het programma berekent automatisch de oplossing voor één cluster tot en met een opgegeven aantal clusters (met een maximum van 40). De variabelen waarop geclusterd wordt kunnen eventueel vooraf worden gestandaardiseerd. Hun aantal is beperkt tot maximaal 25.

HICLUST

HICLUST begint met aan te nemen dat alle subjecten samen (maximaal 200) een cluster vormen. Dan wordt nagegaan of, en eventueel hoe, dit cluster gesplitst kan worden. Deze procedure gaat door totdat er geen significante splitsing van een cluster meer mogelijk is. In de uitvoer van HICLUST worden de clusters in de verschillende stadia gepresenteerd. Bovendien wordt voor elk cluster de zogenaamde Johnson cluster statistic (Johnson, 1967) gegeven.

WARD

De methode van Ward is een agglomeratieve hiërarchische clustertechniek. De methode begint met alle subjecten als losse clusters en gaat vervolgens clusters samenvoegen. Clusters worden zodanig samengenomen dat de som van de varianties binnen de clusters zo weinig mogelijk toeneemt. Het algoritme stopt als alle subjecten samen één cluster vormen. Er kunnen maximaal 300 subjecten worden geclusterd op basis van maximaal 300 variabelen. Van de eerste clusterniveaus wordt een vrij summiere uitvoer gegeven: de som van de varianties binnen de clusters, de grootte van de clusters en welke clusters worden gecombineerd. Van de niveaus met minder dan tien clusters worden ook de labels van de subjecten in de clusters gegeven.

3.4. Modules voor correspondentieanalyse

Voor correspondentieanalyse is een tweetal modules opgenomen, namelijk CORAN en CORRESP. Het belangrijkste verschil tussen CORAN en CORRESP is dat in de eindoplossing van CORAN alleen afstanden tussen rijpunten onderling vergeleken mogen worden. Volgens Carroll, Green en Schaffer (1986 en 1987) mogen bij CORRESP ook de afstanden tussen rijpunten en kolompunten geïnterpreteerd worden. Dit is in veel praktische toepassingen van groot belang. Daarom is het aan te raden CORRESP te gebruiken. De modules CORRESP en CORAN van PC-MDS worden in Fox (1988) vergeleken met een aantal andere programma's voor correspondentieanalyse. Deze vergelijkende test wordt door PC-MDS goed doorstaan.

CORAN

Het algoritme hiervoor is meer traditioneel en is ontwikkeld door Lebart, Morineau en Warwick (1984). Zij behoren tot de zogenaamde 'Franse School' van de grondlegger van correspondentieanalyse Benzecri. Het programma kan een matrix analyseren, waarvan het aantal rijen maal het aantal kolommen niet groter is dan 2000. Een nadeel van deze module is dat de input aan zeer stringente regels moet voldoen. De uitvoer bevat niet de oorspronkelijke frequentiematrix. Het is wel mogelijk een grafiek te maken, waarin zowel de kolompunten als de rijpunten een label hebben.

CORRESP

Het algoritme dat gebruikt wordt in CORRESP is door Carroll, Green en Schaffer beschreven in een tweetal artikelen in het Journal of Marketing Research (1986 en 1987). Zij beweren dat nu ook afstanden tussen punten uit verschillende sets (dus rijpunten met kolompunten) in de eindoplossing vergeleken mogen worden. In het zelfde tijdschrift wordt deze eigenschap van het algoritme door Greenacre (1989) bestreden. CORRESP is vriendelijker in het gebruik dan CORAN. De parameters kunnen ook met menu's in het programma worden ingevoerd. De invoer hoeft niet aan veel eisen te voldoen, en kan ook in 'free format' worden ingelezen. De oorspronkelijk frequenties zijn nu ook opgenomen in de uitvoer. Ook hier wordt een grafiek getekend. Zowel het aantal rijen als aantal kolommen in de CORRESP invoer is beperkt tot 100.

3.5. Modules voor multivariate analyse

Naast de bovengenoemde modules, zijn in PC-MDS een aantal modules opgenomen waarmee multivariate analyses kunnen worden uitgevoerd. Deze analysemogelijkheden zijn: discriminantanalyse (DISCRIM), factoranalyse (FACTOR), automatic interaction detection (AID) en regressieanalyse (REGRESS).

DISCRIM

De module DISCRIM voert een stapsgewijze discriminantanalyse uit. Voor deze module moet een opdrachtenbestand, vergelijkbaar met een SPSS-opdrachtenbestand, worden gemaakt. Er zijn vier criteria beschikbaar op basis waarvan variabelen geselecteerd kunnen worden. Gewone, niet-stepwise, analyses kunnen echter niet met DISCRIM worden uitgevoerd. Er kunnen a priori kansen worden opgegeven voor het classificeren van de

subjecten. De uitvoer van DISCRIM is zeer uitgebreid.

FACTOR

Met behulp van FACTOR kan de gebruiker van PC-MDS een eenvoudige factoranalyse (principale componentenanalyse) uitvoeren. Ook voor FACTOR moet een opdrachtenbestand worden gemaakt. Het is onder andere mogelijk de 'missing values' te definiëren (zowel listwise als pairwise deletion) en een deel van de variabelen te selecteren. Ook kunnen de factorscores worden bewaard. Verder kan men met FACTOR de factoren zowel orthogonaal als scheef roteren. Het programma biedt niet de mogelijkheid andere factoranalyses uit te voeren dan principale componentenanalyse. Dit is een duidelijk nadeel ten opzichte van bijvoorbeeld SPSS/PC. Zowel FACTOR als SPSS/PC geven overigens helaas geen verklaarde variantie per factor na rotatie.

AID

AID (= automatic interaction detection) bestudeert de onderlinge relaties tussen variabelen. Met behulp van technieken uit de variantie-analyse worden de subjecten via een niet symmetrisch vertakkingsproces in subgroepen verdeeld. De subgroepen worden zo gekozen dat de variantie van de waarden voor de afhankelijke variabele binnen subgroepen minimaal is. In AID kunnen maximaal 60 variabelen geanalyseerd worden op basis van gegevens van minimaal 50 subjecten. Meer dan 90 subgroepen kunnen door AID niet gevormd worden. Voor AID dient vooraf een file met controle-commando's gecreëerd te worden, die geheel anders is dan de SPSS-achtige opdrachtenbestanden van DISCRIM, FACTOR, REGRESS en FREQ. Aangezien de controle-file van AID aan veel eisen moet voldoen, loopt men het risico dat er fouten gemaakt worden.

REGRESS

Net als bij de modules DISCRIM en FACTOR moet voor REGRESS een bestand met de opdrachten worden gecreëerd. In REGRESS is alleen stepwise lineaire regressie opgenomen. Het is eenvoudig mogelijk te specificeren welke variabelen afhankelijk zijn en welke onafhankelijk zijn. Bovendien hoeven niet alle variabelen vooraf geselecteerd te worden. Voor verschillende andere varianten van regressieanalyse zal echter moeten worden uitgeweken naar bijvoorbeeld SPSS/PC of ESP.

3.6. Overige modules

Een aantal modules zijn moeilijk onder te brengen in één van de vijf voorgaande subsecties. Daarom worden deze modules, te weten FMATCH, DISTRAN, CASE5 en FREQ, hier in de subsectie 'Overige modules' besproken.

FMATCH

Hiermee kan een procrustesanalyse worden uitgevoerd op basis van het algoritme van Cliff (1966). Het is de bedoeling twee matrices (b.v. configuraties uit KYST) zoveel mogelijk op elkaar te laten lijken. Dit kan gebeuren door één van beide matrices of beide matrices orthogonaal te roteren. Beide mogelijkheden zijn beschikbaar in FMATCH. Naast een figuur met de gerooteerde oplossing worden de rotatie matrices, een goodness-of-fit maatstaf en de correlaties tussen de afstandsvectoren gegeven in de uitvoer. Er zijn echter ook modernere en meer specifieke programma's verkrijgbaar voor procrustesanalyse (bijvoorbeeld Procrustes/PC, zie Kiers, 1990).

DISTRAN

Deze module kan gebruikt worden om de Euclidische afstanden tussen punten te berekenen. Het maximum aantal variabelen/dimensies is 75. Deze variabelen kunnen worden gestandaardiseerd en gewogen indien gewenst. Het volgende kan als uitvoer worden gegeven: de ruwe data, de gestandaardiseerde scores, de getransponeerde datamatrix, de al dan niet gestandaardiseerde afstanden en de gemiddelde afstand tussen alle punten.

CASES

Met behulp van deze module kunnen ééndimensionale Thurstone-schalen worden geconstrueerd. Mogelijk data types als invoer zijn: paarsgewijze vergelijkingen, rangordes of percentages van subjecten dat de stimuli prefereert. Het maximum aantal stimuli is 60.

FREQ

De mogelijkheden van deze module zijn vrij groot. Vele statistieken voor de centrale tendentie en de spreiding kunnen bij de frequentietabellen geleverd worden. Ook voor FREQ moet vooraf een opdrachtenbestand worden gemaakt.

EDIT

De module EDIT is een ASCII tekstverwerker. Hiermee kunnen zowel de databestanden als de opdrachtenbestanden worden gemaakt. De functietoetsen van EDIT komen sterk overeen met die van WordPerfect. Bovendien kan met meerdere teksten simultaan worden gewerkt.

4. Voorbeeld van een analyse met KYST, PROFIT, CORRESP en FMATCH

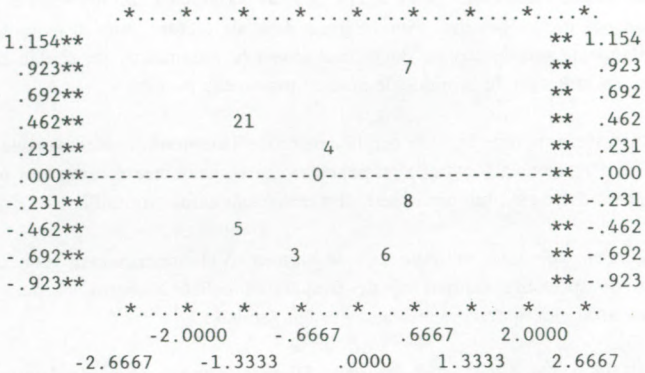
In deze sectie laten we aan de hand van een kort voorbeeld zien wat mogelijk is met PC-MDS. In dit voorbeeld gebruiken we met name die modules van het pakket PC-MDS die naar onze mening het meest interessant zijn.

Stel dat een consument een achttal merken paarsgewijs met elkaar vergeleken heeft. Dit is gebeurd op een negenpuntsschaal, met '1' is sterk gelijk en '9' is zeer ongelijk. Daarnaast heeft deze persoon de acht merken gescoord op een vijftal attributen, achtereenvolgens: goede reclame, mooie verpakking, hoge prijs, goede kwaliteit en traditioneel. Hierbij geldt dat een '5' de maximale score is, die staat voor een merk dat sterk voldoet aan de genoemde eigenschap. De invoerprogramma's zijn opgenomen in de bijlage. Uit deze programma's valt af te lezen hoe de consument in kwestie heeft geantwoord. Opgemerkt dient te worden dat de figuren en tabellen in de oorspronkelijke PC-MDS uitvoer in een aantal gevallen groter zijn dan de figuren en tabellen in dit artikel.

4.1. Voorbeeld van een analyse met KYST en PROFIT

De ongelijkheidsdata kunnen geanalyseerd worden met KYST en/of INDSCAL. Omdat er hier sprake is van slechts één consument wordt KYST gebruikt. Uit het invoerprogramma valt af te lezen wat de antwoorden van de consument zijn. De eerste regel in het invoerprogramma geeft aan dat er 8 stimuli zijn en dat de matrix zowel horizontaal als verticaal 1 maal is opgenomen (alleen relevant bij meerdere respondenten). De tweede regel geeft aan hoe de data gelezen moeten worden. Binnen KYST zijn onder andere de volgende opties uit de menu's gekozen: Stressformule 2, Niet-metrische multidimensional scaling,

Dissimilarities en Primary approach to ties. Dit resulteerde in de onderstaande configuratie, met een stresswaarde van 0.211 in twee dimensies.



Figuur 1. Voorbeeld van een stimulusconfiguratie gegenereerd met KYST.

Om het benoemen van de dimensies te vereenvoudigen zijn de attributen door middel van de module PROFIT in de configuratie opgenomen. De invoerfile voor deze analyse staat in de bijlage. De twee belangrijkste tabellen uit de uitvoer van PROFIT zijn hier onder opgenomen.

TABLE 1. THE MAXIMUM CORRELATION BETWEEN THE PROPERTY AND THE PROJECTIONS ON FITTED VECTOR

	RHO	PROPERTY
1	.5555	goede reclame
2	.5965	mooie verpakking
3	.7695	hoge prijs
4	.9806	goede kwaliteit
5	.4388	traditioneel

Tabel 2. Voorbeeld van uitvoertabel 1 van PROFIT.

Uit deze tabel valt af te lezen welke attributen goed kunnen worden weergegeven als vectoren in de configuratie. Met name de eigenschap 'traditioneel' kan goed worden weergegeven in de stimulusconfiguratie.

TABLE 2. DIRECTION COSINES OF FITTED VECTORS
IN NORMALIZED SPACE

VECTOR	DIMENSION	
	1	2
1	.0287	-.9996
2	-.7253	-.6885
3	-.7914	.6113
4	-.9088	.4172
5	.7262	.6875

Tabel 3. Voorbeeld van uitvoertabel 2 van PROFIT.

Met behulp van deze tabel kunnen de dimensies geïnterpreteerd worden. De eerste dimensie lijkt te benoemen als 'kwaliteit' en de tweede dimensie als 'reclame'. De overige drie eigenschappen liggen tussen de dimensies in.

4.2. Voorbeeld van een analyse met CORRESP

De scores van de merken op de attributen kunnen ook met CORRESP geanalyseerd worden. De gelijkenisdata worden nu niet gebruikt. De configuratie bevat in dit geval zowel de merken als de attributen weergegeven als punten. Voor de data en het invoerprogramma wordt verwezen naar de bijlage. Ook CORRESP geeft een figuur als uitvoer, die is hier echter niet weergegeven. Onderstaande tabel vat de belangrijkste uitkomsten van CORRESP samen.

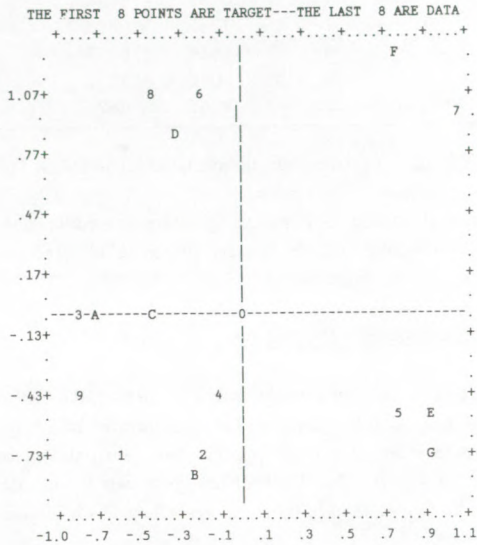
EIGENVALUES OF THE CHI-SQUARE DISTANCE MATRIX				
	1	2	3	4
VALUES	.1054	.0425	.0115	.0030
% OF TOTAL	64.8828	26.1704	7.0994	1.8474
CUMULATIVE %	64.8828	91.0532	98.1526	100.0000

Tabel 4. Voorbeeld van een tabel met eigen waarden en verklaarde varianties van CORRESP.

Het blijkt dat twee dimensies voldoende zijn om meer dan 90 % van de variantie te verklaren.

4.2. Voorbeeld van een analyse met FMATCH

Er zijn nu twee configuraties van de acht merken beschikbaar. In hoeverre komen deze overeen? Dit kan worden nagegaan met behulp van FMATCH. De configuraties worden zo geroteerd dat ze zo veel mogelijk op elkaar lijken. Dit resulteert in de onderstaande figuur uit de FMATCH-uitvoer.



Figuur 2. Voorbeeld van een FMATCH configuratie.

Bovendien worden nog een aantal fitmaatstaven gegeven voor de gevonden oplossing. Deze zijn weergegeven in de onderstaande tabel.

THE GOODNESS OF FIT IS .610 -1. IS WORST, +1. IS PERFECT		
CORRELATION MATRIX		
ROW: 1	.939	.000
ROW: 2	.000	.522
THE CORRELATION OF THE DISTANCE VECTORS IS .637		

Tabel 5. Fitmaatstaven in de FMATCH uitvoer.

De configuraties van KYST en CORRESP lijken niet erg sterk op elkaar. Met name op de tweede dimensie van figuur 2 zijn de merken verschillend gepositioneerd.

5. De handleiding

De opzet van de PC-MDS handleiding is uitvoerig (309 bladzijden) en overzichtelijk. Na de benodigde algemene informatie over het pakket, wordt uitgelegd hoe de invoerbestanden van DISCRIM, FACTOR, REGRESS en FREQ moeten worden gemaakt. Daarna worden de afzonderlijke modules besproken. De beschrijving van elke module bestaat uit drie delen. Als eerste een theoretische beschrijving en als tweede een overzicht van de mogelijkheden en de daarvoor te kiezen parameterwaarden. Daarna volgt een bespreking van een analyse voorbeeld aan de hand van de invoer en de uitvoer. Bij een groot aantal modules wordt in de verbrede kantlijn van de uitvoer een toelichting gegeven op de betreffende figuren, tabellen en statistieken. Vooral dit derde deel is zeer nuttig. Het aantal pagina's dat wordt besteed aan de uitleg van een module varieert nogal sterk. De beschrijving van KYST, bijvoorbeeld, neemt 25 pagina's in beslag. De handleiding bij de multidimensional scaling modules is dan ook goed te noemen. Ook de bespreking van zowel de theorie als het gebruik van de modules waarbij een opdrachtenbestand moet worden gemaakt (DISCRIM, FACTOR, REGRESS en FREQ) is zeer uitvoerig. Een aantal andere modules worden vrij kort besproken. Zo is de bespreking van de naar onze mening zeer bruikbare module CORRESP slechts 11 pagina's lang. Wel wordt in dergelijke gevallen verwezen naar literatuur over de theorie van de betreffende analyse.

Een duidelijk minpunt van de handleiding is de (on-)nauwkeurigheid. Er zijn vele fouten in de handleiding geslopen. Met name de parameteropties verschillen in een aantal gevallen tussen de handleiding en de eigenlijke programmatuur. Bovendien verschillen de besprekingen van de modules onderling in niveau en in structuur.

6. Relatie met het boek van Green, Carmone en Smith

Een vereenvoudigde versie van het PC-MDS pakket is bijgesloten in het boek van Green, Carmone en Smith (1989) (voor een recensie, zie Vriens, 1989). In deze versie zijn echter slechts zeven modules opgenomen, namelijk KYST, INDSCAL, PREFMAP, PROFIT, MDPREF, CORRESP en CLUSTER. Deze modules zijn bovendien sterk gereduceerd wat betreft de invoermogelijkheden. De verkorte versie neemt dan ook slechts 2 floppy's van elk 360 KB in beslag. De delen van de handleiding die behoren bij de bovengenoemde zeven modules zijn opgenomen in Green, Carmone en Smith (1989). Dit boek met bijbehorende floppy's kost ongeveer f 120,-.

7. Verschil PC-MDS versie 5.1 met versie 4.1

De nieuwe versie van PC-MDS ondervangt een aantal duidelijke minpunten van de voorgaande versies. Ten eerste is de module EDIT toegevoegd, zodat voor het maken van de opdracht- en databestanden niet meer hoeft te worden uitgeweken naar een ander pakket. Bovendien is de bruikbaarheid van de op zich minder interessante modules DISCRIM, FACTOR, REGRESS en FREQ verbeterd door te werken met een SPSS-achtig

opdrachtenbestand. Dit maakt dat het pakket PC-MDS voor een onderzoeker die meerdere (met name ook eenvoudiger) analysetechnieken wil gebruiken interessanter is geworden. Als derde vernieuwing valt op te merken dat na de modules die resulteren in een stimulusconfiguratie (de modules voor multidimensional scaling en correspondentieanalyse) een driedimensionale figuur op het scherm wordt weergegeven. Deze figuur kan met de pijltjestoetsen worden gerooteerd. Bovendien kan de figuur op schijf worden bewaard of direct worden geprint. Ten vierde zijn door het gehele pakket heen de beperkingen gewijzigd. Waar nodig is bijvoorbeeld het maximum aantal stimuli uitgebreid. Hier en daar is het pakket verder beperkt. Dit is echter gedaan op plaatsen waar de oudere versies van het pakket, veelal onnodig, grote invoer accepteerde. Ten vijfde is het pakket gebruikersvriendelijker geworden. Het is niet meer zo dat het onjuist specificeren van een parameter of het onjuist opgeven een naam van een bestand resulteert in het direct en onvrijwillig verlaten van het programma. Tenslotte is de handleiding aangepast. Het lettertype en mede daardoor het gehele uiterlijk van de nieuwe handleiding is fraaier dan in de oude handleiding. Bovendien is de uitleg van de theorie en van de opties, met name bij de modules DISCRIM, FACTOR, REGRESS en FREQ, uitgebreid en verbeterd.

8. Algemene evaluatie van PC-MDS

Het pakket bevat zoals reeds aangegeven een verzameling modules. Deze modules zijn echter samengevoegd zonder dat er een homogeen pakket van gemaakt is waarmee op gebruikersvriendelijke wijze gewerkt kan worden. Het is niet mogelijk een databestand te analyseren met meerdere modules, zonder een groot aantal aanpassingen te maken. Het pakket zal daarom waarschijnlijk niet bij een groot publiek aanslaan. Toch kan het voor onderzoekers een aantal interessante mogelijkheden bieden. Het multidimensional scaling gedeelte is ons inziens het meest aantrekkelijk, omdat het hier de bekende programma's betreft (zoals KYST, PREFMAP, etc.) die niet elders in een statistisch pakket opgenomen zijn. Daarnaast kan het pakket interessant zijn voor correspondentieanalyse, met name vanwege de opname van het Carroll, Green en Schaffer algoritme (CORRESP). Ook conjuncte analyse kan aantrekkelijk zijn vanwege de modules CONJOINT en TRADEOFF. De modules FMATCH (procrustesanalyse) en AID (automatic interaction detection) zijn goed te gebruiken. De modules voor clusteranalyse, discriminantanalyse, factoranalyse, regressieanalyse en frequentietabellen zijn minder interessant voor de gebruiker, omdat deze technieken uitgebreider in de meer gebruikersvriendelijke pakketten zoals SPSS/PC en SAS zijn opgenomen. De bruikbaarheid van de modules DISCRIM, FACTOR, REGRESS en FREQ is echter wel aanzienlijk verbeterd ten opzichte van eerdere versies van PC-MDS.

Literatuur

- Carroll, J.D. en J.J. Chang (1970), Analysis of individual differences in Multidimensional Scaling via an n-way generalization of 'Eckart-Young' decomposition, *Psychometrika*, 35, 283-319.
- Carroll, J.D., P.E. Green en C.M. Schaffer (1986), Interpoint distance comparisons in

- Correspondence Analysis, Journal of Marketing Research, 23, 271-280.
- Carroll, J.D., P.E. Green en C.M. Schaffer (1987), Comparing interpoint distances in Correspondence Analysis: a clarification, Journal of Marketing Research, 24, 445-450.
- Cliff, N. (1966), Orthogonal rotation to congruence, Psychometrika, 31, 33-42.
- Deth, J.W. van (1979), PROFIT: a program for FITting PROPERTIES in a multidimensional space using linear or nonlinear regression procedures, TC-publicatie nr. 92, Universiteit van Amsterdam.
- Fox, R.J. (1988), Review of Correspondence Analysis PC software packages, Journal of Marketing Research, 25, 414-417.
- Green, P.E., F.J. Carmone en S.M. Smith (1989), Multidimensional Scaling: concepts and applications. Boston: Allyn and Bacon.
- Green, P.E. en V. Srinivasan (1978), Conjoint Analysis in consumer research: Issues and outlook, Journal of Consumer Research, 5, 103-123.
- Greenacre, M.J. (1989), The Carroll-Green-Schaffer Scaling in Correspondence Analysis: a theoretical and empirical appraisal, Journal of Marketing Research, 26, 358-365.
- Johnson, S.C. (1967), Hierarchical clustering schemes, Psychometrika, 32, 241-254.
- Kiers, H.A.L. (1990), PROCUSTES-PC v2.0: een programma voor gegeneraliseerde Procrustes-analyse, Kwantitatieve Methoden, nr. 33, 177-188.
- Kruskal, J.B. (1964), Multidimensional Scaling by optimizing goodness of fit to a nonmetric hypothesis, Psychometrika, 29, 1-27. (a)
- Kruskal, J.B. (1964), Nonmetric Multidimensional Scaling: a numeric method, Psychometrika, 29, 115-129. (b)
- Kruskal, J.B. en F.J. Carmone (1969), MONANOVA: a FORTRAN-IV program for monotone analysis of variance, Behavioral Science, 14, 165-166.
- Lebart, L., A. Morineau en K.M. Warwick (1984), Multiaattribute descriptive statistical analysis: Correspondence Analysis and related techniques for large matrices. New York: Wiley.
- Nehls, L., B. Seamon en D. Montgomery (1974), ETROFF: a basic program for Trade-off analysis, Report number 76-1051, Marketing Science Institute.
- Punj, G. en D.W. Stewart (1983), Cluster analysis in marketing research: review and suggestions for application, Journal of Marketing Research, 20, 134-148
- Schuur, W.H. van (1976), Preference Mapping: an introduction to Carroll and Chang's program PREFMAP, interne publikatie, Rijksuniversiteit Groningen.
- Vriens, M. (1989), Boekbespreking van Green, Carmone en Smith (1989), Kwantitatieve Methoden, nr. 33, 208-210.
- Vriens, M. en D.R. Wittink (1990), Conjuncte Analyse in het marktonderzoek, in: Jaarboek van de Nederlandse Vereniging van Marktonderzoekers '90-'91, Haarlem: De Vriesebosch, 215-246.
- Vriens, M. en T. Wilms (1990), ACA: Adaptieve Conjuncte Analyse, Een software pakket voor Conjuncte Analyse, Kwantitatieve Methoden, nr. 37, 139-156.
- Wittink, D.R. en P. Cattin (1989), Commercial use of Conjoint Analysis: An update, Journal of Marketing Research, 26, 91-96.
- Young, F.W. en W.S. Torgerson (1967), TORSCA: a FORTRAN-IV program for Shepard-Kruskal Multidimensional Scaling Analysis, Behavioral Science, 12, 498.

Bijlage: Invoerbestanden van de voorbeeld analyses

KYST	PROFIT	CORRESP	FMATCH
8 1 1	1 8 2 5 0 0 1 0.0	8 5	8 2 2 1 2
(8f1.0)	(8f7.3)	(5f2.0)	(3x,2f6.3)
1	-.931 .430	4 5 5 4 1	(3x,2f6.3)
43	-.937 .426	3 4 4 4 2	c1 -.942 -.923
533	-.295 -.588	5 5 3 3 1	c2 -.255-1.029
5453	.069 .139	3 3 4 3 2	c3 -1.184 .211
67426	-.958 -.417	2 2 3 4 4	c4 -.149 -.592
877565	.812 -.671	4 4 2 1 3	c5 1.293 -.898
7654713	1.121 .834	1 1 2 2 5	c6 -.043 1.790
	1.120 -.155	5 2 3 1 2	c7 2.868 .360
	(8f2.0)	merk 1	c8 -.440 1.907
	goede reclame	merk 2	k1 -.931 .430
	4 3 5 3 2 4 1 5	merk 3	k2 -.937 .426
	mooie verpakking	merk 4	k3 -.295 -.588
	5 4 5 3 2 4 1 2	merk 5	k4 .069 .139
	hoge prijs	merk 6	k5 -.958 -.417
	5 4 3 4 3 2 2 3	merk 7	k6 .812 -.671
	goede kwaliteit	merk 8	k7 1.121 .834
	4 4 3 3 4 1 2 1	goede reclame	k8 1.120 -.155
	traditioneel	mooie verpakking	
	1 2 1 2 4 3 5 2	hoge prijs	
		goede kwaliteit	
		traditioneel	