

Commentaar op
"Procrustes-PC V2.0:
Een programma voor gegeneraliseerde Procrustes-analyse"
door Henk A.L. Kiers

Stef van Buuren
Vakgroep Psychonomie
Rijksuniversiteit Utrecht

Garmt Dijksterhuis
OP&P, Postbus 14167, 3508 SG, Utrecht

Introductie

Allereerst willen wij Kiers (1990) bedanken voor de uitgebreide aandacht die hij geschonken heeft aan de bespreking. Terecht merkt hij een aantal onvolkomenheden op zoals het crashen en ander onverwacht gedrag van het programma, en inconsistenties en omissies in de handleiding. Zijn notities kunnen slechts leiden tot een verbetering van het programma en van de documentatie.

Met zijn conclusie "niet direkt aanbevelenswaardig" zijn wij het echter niet eens. Hij baseert zich op drie argumenten: gebrek aan keuze van het criterium, onmogelijkheid van random startconfiguraties en het gebrek aan gebruikersvriendelijkheid. Hieronder worden kort een aantal keuzes toegelicht die bij het ontwerp van het programma gemaakt zijn.

Optimalisatie criterium

Het is tot op heden allerminst duidelijk welk van de door Peay (1988) genoemde criteria in de praktijk de voorkeur geniet. Het is voor de gebruiker dan ook niet bijster interessant om te kunnen kiezen uit verschillende minimalisatie criteria wanneer hem niet duidelijk kan worden gemaakt wanneer voor welke variant te kiezen. Mensen die überhaupt in staat zijn een gefundeerd oordeel te geven en de verschillen tussen de criteria te begrijpen zijn schaars en zijn zeker niet in de toegepaste sfeer te vinden. Het inbouwen van andere criteria kan weliswaar vrij eenvoudig geschieden, maar het resultaat zal voor de meeste gebruikers slechts een extra en weinig gebruikte optie zijn.

De verschillende criteria leiden ons inziens slechts tot gradueel verschillende oplossingen. Vooral de consensusruimtes lijken sterk op elkaar. Misschien is juist dit aanleiding om meerdere criteria te bieden zodat dit soort overeenkomsten beter onderzocht kan worden. Voor dit doel is het programma echter niet opgezet. Wel om op eenvoudige wijze gegevens te kunnen analyseren. Een praktisch argument om vooralsnog het consensus criterium boven het kleinste kwadraten criterium te kiezen is dat voor het eerste een efficiënter algoritme beschikbaar is. Peay noemt in dit verband een faktor 5. Daarnaast heeft het criterium de aantrekkelijke eigenschap dat de consensus variantie zo groot mogelijk wordt. Dit is vooral van belang bij

dimensie-redukcie.

In de Appendix beschrijft Kiers een voorbeeld dat suggereert dat in dit geval het kleinste kwadraten criterium "beter" is dan het consensus criterium (overigens mist H1 een minteken). Met evenveel recht kan men echter stellen dat hier het consensus criterium "beter" is, aangezien de consensusruimte een grotere proportie variantie verklaart. De laatste oplossing geeft dus een completere beschrijving van de data. Het voorbeeld laat slechts zien dat verschillende criteria leiden tot verschillende oplossingen. Dit is niets nieuws. Dat het kleinste kwadraten verlies gelijk is aan nul doet niet echt terzake.

Gebruikersvriendelijkheid

Voor wie weleens een Procrustes programma op een mainframe heeft gedraaid zal het PC programma een verademing zijn. De gangbare mainframe programma's zijn voor het merendeel echte rekenmonsters. Ze zijn matig tot zeer slecht te bedienen, hebben bijzonder bescheiden in- en uitvoeropties, zijn zelden van een deugdelijke handleiding voorzien en worden derhalve slechts door een handvol experts gebruikt. In dit licht bezien is Procrustes-PC een flinke vooruitgang: het kan op iedere MS-DOS machine draaien, het biedt bijvoorbeeld de mogelijkheid om van elk resultaat een plot te maken, het te listen of apart te bewaren, de aansturing wordt verzorgd door begrijpelijke SPSS-achtige commando's en zo is er nog een serie verbeteringen ten opzichte van de zware broeders. Graag zouden we dan ook gezien hebben dat Kiers het programma op dit punt met de concurrentie zou hebben vergeleken.

Uiteraard kan het altijd mooier en beter, maar aan de suggestie om het programma interactief te maken kleeft wel een nadeel. Bij een rekenintensieve toepassing als Procrustes analyse zal de gebruiker onnodig lang moet wachten op de resultaten. Een batch-job kan in principe in de achtergrond gedraaid worden en maakt de machine vrij voor andere werkzaamheden. Daarnaast is het bouwen van een goed interactief gebruikersinterface bepaald geen sinecure en vereist dit veel toewijding, geduld en geld, zeker als men daarbij wil dat de berekeningen in de achtergrond worden uitgevoerd.

Precisie

De waarschuwing het convergentie criterium niet te strak te zetten is niet voor niets in de handleiding opgenomen. Het algoritme divergeert bij zeer kleine convergentiewaarden. Dit wordt veroorzaakt doordat de data en de tussenresultaten worden opgeslagen met een precisie van zes decimalen. Door deze precisie te verdubbelen wordt het mogelijk numeriek stabielere berekeningen uit te voeren. Convergentie kan dan gegarandeerd worden tot ongeveer 12 decimalen. In Procrustes-PC is echter gekozen om in enkele precisie te werken op grond van twee redenen. Ten eerste wordt het programma hierdoor efficiënter, vooral wat betreft geheugengebruik. Ten tweede zit men doorgaans niet te wachten op al die extra decimalen en spelen deze nauwelijks een rol bij de interpretatie van de analyse. Dit laatste is uiteraard afhankelijk van de schaal van de gegevens. Deze schaal kan door de onderzoeker echter

eenvoudig gemanipuleerd worden, zodat de keuze voor een te streng convergentie-criterium in voorkomende gevallen omzeild kan worden.

Conclusie

Het programma is bedoeld om grotere groepen onderzoekers de mogelijkheid te geven Procrustes analyse toe te passen. De nadruk ligt op het toegankelijk maken van de techniek voor niet-specialisten. In eerste instantie is daarom niet gekozen om alle mogelijkheden die door Peay (1988) behandeld worden te implementeren. Het programma beperkt zich tot een belangrijke subset daarvan. Overigens dient hierbij opgemerkt te worden dat er voor zover wij hebben kunnen nagaan geen enkel ander programma verkrijgbaar is dat in staat is zowel met een ongelijk aantal variabelen per matrix te werken alsmede deze matrices te schalen. In dit opzicht is Procrustes-PC uniek en kan het meer dan menig ander rotatie programma. Een aanvraag voor het programma van Peay leverde geen reactie op.

Het programma bezit geen gebruikersinterface met menu's en windows. Wel is het huidige commando gestuurde programma een forse verbetering ten opzichte van de bestaande Procrustes programmatuur. Daarnaast heeft het programma tal van nuttige extra's in de vorm van plots, data-selektie, foutendiagnose, variantie-tabellen en meer.

Het huidige programma blijkt een aantal fouten niet correct af te vangen. Dit is vanzelfsprekend niet te rechtvaardigen. Deze problemen zijn met de eerstvolgende release verholpen. Daarnaast doet Kiers enkele nuttige suggesties (uitgebreidere seed, rotatie naar simple structure, ontkoppeling normalisatie-translatie, inproduct algoritme, grotere data opslag). Hiervan stonden er reeds enige op de lijst. Afgezien van het crashen kan echter ook de huidige versie de vergelijking met elk soortgelijk programma met gemak doorstaan.

Referenties

- Kiers, H.A.L. (1990) Procrustes-PC V2.0: Een programma voor gegeneraliseerde Procrustes-analyse. *Kwantitatieve Methoden*, 33, 177-188
- Peay, E.R. (1988) Multidimensional rotation and scaling of configurations to optimal agreement. *Psychometrika*, 53, 2, 199-208