
References

- Borg, I. & Lingoes, J. (1987). *Multidimensional similarity structure analysis*. New York: Springer Verlag.
- Borg, I. (1979). Geometric representation of individual differences. In Lingoes et al. (Eds), *Geometric representations of relational data* (pp. 609 - 656). Ann Arbor: Mathesis Press.
- Carroll, J.D. & Chang, J.J. (1970). Analysis of individual differences in multidimensional scaling via an N-way generalization of Eckart-Young decomposition. *Psychometrika*, **35**, 283 - 319.
- Carroll, J.D. & Chang, J.J. (1972). *IDIOSCAL (Individual Differences In Orientation SCALing): A generalization of INDSCAL allowing IDIOSyncratic reference systems as well as an analytic approximation to INDSCAL*. Paper presented at meetings of the Psychometric Society, Princeton, N.J.
- Cliff, N. (1966). Orthogonal rotation to congruence. *Psychometrika*, **31**, 33 - 42.
- Coxon, A.P.M. (1982). *The user's guide to multidimensional scaling*. London: Heinemann Educational Books.
- Davies, P.M. & Coxon, A.P.M. (1983). *The MDS(X) series of multidimensional scaling programs*. Report No 55, University of Edinburgh.
- Davison, M.L. (1983). *Multidimensional scaling*. New York: Wiley.
- De Leeuw, J. & Meulman, J. (1986). A special jackknife for multidimensional scaling. *Journal of Classification*, **3**, 97-112.
- Everitt, B.S. & Gower, J.C. (1981). Plotting the optimum positions of an array of cortical electrical phosphenes. In Barnett (Ed), *Interpreting multivariate data*. (pp. 279 - 287). Chichester: Wiley.
- Gifi, A. (1990). *Nonlinear multivariate analysis*. New York: Wiley.
- Gower, J.C. (1975). Generalized Procrustes analysis. *Psychometrika*, **40**, 33 -51.
- Green, B.F. (1952). The orthogonal approximation of an oblique structure in factor analysis. *Psychometrika*, **17**, 429 - 440.
- Green, P.E., Carmone, F.J., & Smith, S.M. (1989). *Multidimensional scaling*. Boston: Allyn and Bacon.
- Horan, C.B. (1969). Multidimensional scaling: Combining observations when individuals have different perceptual structures. *Psychometrika*, **34**, 139 - 165.
- Kristof, W. (1970). A theorem on the trace of certain matrix products and some applications. *Journal of Mathematical Psychology*, **7**, 515 - 530.
- Kristof, W. & Wingersky, B. (1971). Generalization of the orthogonal Procrustes rotation procedure to more than two matrices. *Proceedings of the 79th Annual Convention of the American Psychological Association*, **6**, 89 - 90.

- Lawson, C.L. & Hanson, R.J. (1974). *Solving least squares problems*. Englewood Cliffs, NJ: Prentice Hall.
- Langeheine, R. (1982). Statistical evaluation of measures of fit in the Lingoes-Borg Procrustean individual differences scaling. *Psychometrika*, **47**, 427 - 442.
- Lingoes, J.C. & Borg, I. (1978). A direct approach to individual differences scaling using increasingly complex transformations. *Psychometrika*, **43**, 491 - 519.
- Lingoes, J.C. & Schönemann, P.H. (1974). Alternative measures of fit for the Schönemann-Carroll matrix fitting algorithm. *Psychometrika*, **39**, 423 - 427.
- Magnus, J.R. & Neudecker, H. (1988). *Matrix differential calculus with applications in statistics and econometrics*. Chichester: Wiley.
- Mooijaart, A. & Commandeur, J.J.F. (in press). A general solution of the weighted orthonormal Procrustes problem. *Psychometrika*.
- Peay, E.R. (1988). Multidimensional rotation and scaling of configurations to optimal agreement. *Psychometrika*, **53**, 199 - 208.
- Schönemann, P.H. (1966). A generalized solution of the orthogonal Procrustes problem. *Psychometrika*, **31**, 1 - 10.
- Schönemann, P.H. & Carroll, R.M. (1970). Fitting one matrix to another under choice of a central dilation and a rigid motion. *Psychometrika*, **35**, 245 - 255.
- Tatsuoka, M.M. (1988). *Multivariate analysis: Techniques for educational and psychological research* (2nd ed.). New York: Macmillan.
- Ten Berge, J.M.F. (1977). Orthogonal Procrustes rotation for two or more matrices. *Psychometrika*, **42**, 267 - 276.
- Ten Berge, J.M.F. (1983). A generalization of Kristof's theorem on the trace of certain matrix products. *Psychometrika*, **48**, 519-523.
- Ten Berge, J.M.F. (1988). Generalized approaches to the Maxbet problem and the Maxdiff problem, with applications to canonical correlations. *Psychometrika*, **53**, 487 - 494.
- Ten Berge, J.M.F. & Knol, D.L. (1984). Orthogonal rotations to maximal agreement for two or more matrices of different column orders. *Psychometrika*, **49**, 49 - 55.
- Van de Geer, J.P. (1984). Linear relations among k sets of variables. *Psychometrika*, **49**, 79 - 94.
- Van de Geer, J. P. (1986). *Introduction to linear multivariate analysis* (2 Vols). University of Leiden: DSWO Press.
- Young, F.W. & Hamer, R.M. (1987). *Multidimensional scaling: history, theory, and applications*. NJ: Lawrence Erlbaum Associates.
- Van der Kloot, W.A. (1978). Matching five representations of the implicit theory of personality. Unpublished manuscript.

Samenvatting

In de afgelopen decennia is er een groot aantal data analyse technieken ontwikkeld die het letterlijk mogelijk maken om 'de werkelijkheid in kaart te brengen'. Deze veelal exploratieve technieken zijn er op gericht om de relaties tussen objecten te onderzoeken door deze als punten af te beelden in een m-dimensionale ruimte. Meestal wordt hierbij de voorkeur gegeven aan een afbeelding in zo min mogelijk dimensies. Zo'n afbeelding van objecten als punten in m dimensies, ook wel een *configuratie* genoemd, is nuttig omdat deze inzicht kan geven in de onderlinge relaties van de objecten. Bovendien biedt de ordening van de objecten op de afzonderlijke dimensies de mogelijkheid om aan die dimensies een interpretatie te geven.

Mede als gevolg van deze ontwikkelingen is de behoefte ontstaan om configuraties *onderling* met elkaar te vergelijken. In allerlei situaties is het van belang om vast te kunnen stellen in hoeverre configuraties identiek zijn, en zo nee, of er een systematisch patroon in de verschillen zit. Een voorbeeld is de situatie waarbij een aantal onderzoekers gegevens verzameld hebben over eenzelfde onderwerp, en waarbij iedere onderzoeker zijn of haar gegevens heeft omgezet in een configuratie. Het wordt dan interessant om een uitspraak te kunnen doen over de mate van overeenkomst (of: *match*) tussen de configuraties.

Het lijkt misschien alsof we in dit soort gevallen gemakkelijk zouden kunnen volstaan met het 'op het oog' vergelijken van de configuraties om vervolgens op grond daarvan te besluiten of ze al dan niet overeenkomen. Maar deze aanpak is weinig systematisch, en heeft bovendien alleen kans van slagen bij één- of tweedimensionale configuraties, d.w.z. bij afbeeldingen op de rechte lijn of in het platte vlak. Zodra we te maken hebben met afbeeldingen van objecten in drie of meer dimensies wordt zo'n visuele vergelijking al vrijwel onuitvoerbaar.

Van de eerder in de literatuur voorgestelde methoden voor onderzoek naar de match tussen configuraties biedt het zg. PINDIS programma (Procustean INdividual Differences Scaling) van Lingoes en Borg (1978) de meest uitgebreide mogelijkheden. Niet alleen kunnen met dit programma de overeenkomsten tussen twee of meer configuraties worden onderzocht aan de hand van maar liefst vijf verschillende modellen, maar deze modellen zijn bovendien hiërarchisch opgebouwd, zodat nagegaan kan worden of ingewikkelder modellen iets toe te voegen hebben aan meer simpele modellen.

Aan het PINDIS programma kleven echter een aantal belangrijke bezwaren. Met name de algoritmen die in dit programma gebruikt worden zijn duidelijk voor verbetering vatbaar. Het belangrijkste oogmerk van dit boek is dan ook om nieuwe convergerende kleinste kwadraten algoritmen aan te dragen voor het schatten van de parameters in de respectievelijke modellen. Bovendien worden de eerder voorgestelde modellen zodanig uitgebreid dat de overeenkomst tussen twee of meer configuraties ook onderzocht kan worden als de configuraties informatie bevatten over *verschillende aantallen objecten*. De in dit boek voorgestelde algoritmen zijn geïmplementeerd in het zg. MATCHALS programma (MATCHing configurations by Alternating Least Squares).

In hoofdstuk 1 worden de vijf modellen behorend bij het PINDIS programma geïntroduceerd, en wordt een gedetailleerd overzicht gegeven van de bezwaren die tegen deze modellen kunnen worden aangevoerd, alsmede tegen de algoritmen voor het schatten van de betreffende modelparameters. Alternatieve modellen worden gepresenteerd, waarbij rekening wordt gehouden met de mogelijkheid van ontbrekende gegevens.

In hoofdstuk 2 wordt de overeenkomst tussen twee of meer configuraties onderzocht onder transformaties die de relatieve afstanden tussen de object punten onveranderd laten. Deze techniek, die ook wel bekend is onder de naam gegeneraliseerde Procrustes analyse (GPA), wordt uitgebreid naar het geval dat de configuraties ongelijke aantallen rijen hebben. Aangetoond wordt dat het bij onvolledige configuraties essentieel is om te corrigeren voor verschillen in grootte tussen de configuraties. Door eliminatie van twee van de vier sets te schatten parameters ontstaat een efficiënt alternerend algoritme voor het schatten van de overgebleven twee transformaties. Het theoretische gedeelte van het hoofdstuk wordt afgesloten met een decompositie van kwadratensommen, en met een bespreking van de overeenkomsten tussen de criteria MAXBET, MAXDIFF, en MAXNEAR en het match criterium in GPA voor incomplete configuraties. Tenslotte worden de resultaten besproken van de analyse van twee getallenvoorbeelden met het in hoofdstuk 2 ontwikkelde algoritme.

In hoofdstuk 3 worden nieuwe convergerende algoritmen ontwikkeld voor het schatten van de parameters in twee zg. *dimensie gewogen* modellen. Deze modellen hebben gemeen dat de (optimaal gekozen) assen van een (vrij te schatten) groepsconfiguratie verschillend mogen worden opgerekt. In het ene model

(DIMFREE genoemd) zijn deze assen identiek voor alle individuele configuraties, terwijl het andere model (DIMIDIO genoemd) de vrijheid biedt om voor iedere individuele configuratie een eigen optimale assenstand in de groepsconfiguratie te bepalen. Voor beide modellen wordt een ontbinding van kwadratensommen gepresenteerd, en de in hoofdstuk 3 ontwikkelde algoritmen worden uitgetest op een aantal geconstrueerde data sets.

In hoofdstuk 4 worden convergerende algoritmen opgezet voor het bepalen van de parameters in twee *zg. stimulus gewogen* modellen. De overeenkomst tussen deze twee modellen is dat de object punten van een (vrij te schatten) groepsconfiguratie verbonden worden met de oorsprong, en dat de hieruit resulterende vectoren verschillend mogen worden verlengd of ingekort. In het ene model (STIMFREE genaamd) is de oorsprong waarmee de object punten van de groepsruimte worden verbonden identiek voor alle individuele configuraties. In het andere model (STIMIDIO genoemd) wordt voor iedere individuele configuratie een eigen optimale oorsprong geschat. Het hoofdstuk wordt weer afgesloten met een decompositie van kwadratensommen voor de twee stimulus gewogen modellen, waarna de resultaten van het STIMFREE en het STIMIDIO algoritme worden geïllustreerd met een aantal geconstrueerde getalenvoorbeelden.

In hoofdstuk 5 tenslotte worden de belangrijkste resultaten uit de voorafgaande hoofdstukken samengevat, en een aantal suggesties gegeven voor verder onderzoek.

MATCHING CONFIGURATIONS

For many multidimensional data analysis methods, such as multidimensional scaling, principal components analysis and correspondence analysis, the major result consists of a configuration of points, that is, a geometrical representation of the relationships between the observational units, or objects. The growing use of these methods has led to an increased interest in the possibility of comparing (or: matching) configurations. Whenever one analyses the same data with different multidimensional techniques, or when one is interested in the question whether the configurations obtained in a number of studies in some common field of research are identical, the problem arises how to evaluate the similarity of the outcomes.

This book starts out from the PINDIS framework for the matching of two or more configurations. It discusses five models in detail, providing convergent alternating least squares algorithms for the estimation of the corresponding model parameters. The first and most parsimonious model investigates the match under transformations that leave the relative distances between the object points intact: Generalized Procrustes analysis. In the second and third model a transformation called dimension weighting is introduced, where dimensions are allowed to be stretched or shrunk differently. In terms of interpretation these cases are related to the well-known INDSCAL and IDIOSCAL models. The fourth and fifth model are characterized by a so-called stimulus weighting transformation, where the centrality of the object points is differentially increased or decreased. These five models form two hierarchies, within each of which it is possible to assess whether more complex transformations have anything new to add over the simpler ones.

When the configurations have been obtained from different studies, it often occurs that the number of objects varies from configuration to configuration. In this case some or all configurations contain missing points. An important new feature of the five algorithms presented in this book is that they can also be used to investigate the match between incomplete configurations.

DSWO PRESS

ISBN 90-6695-052-8