A program to determine the optimal rank order of pairwise similarities from incomplete triads data: Profit

MARK MOODY and EDWARD D. FAHRMEIER

University of Maryland Baltimore County
Baltimore, Maryland 21228

Description. Nonmetric scaling techniques (Kruskal, 1964; Shepard, 1962a, b; Shepard, 1966) as embodied in the TORSCA (Young & Torgerson, 1967) or PARSCAL (Johnson, 1973) programs produce multidimensional scaling solutions from rank orders of pairwise similarities between N points. The present program produces a rank order for $\binom{N}{2}$ pairs of stimuli for data gathered by Torgerson's method of triadic combinations (Torgerson, 1958) or, more commonly, the method of incomplete triads. Coombs (1964, pp. 352-359) suggests a method of triangular analysis to arrange these triad judgments into a single rank order. We have supplemented Coombs' basic algorithm with a set of routines to transform triangular matrices to produce the best rank orders in terms of the fewest number of inconsistencies.

With sophisticated subjects, the rank orders needed for the nonmetric scaling programs can be obtained directly by asking the subject to rank order all pairs of stimuli in order of similarity. The method of incomplete triads is particularly appropriate for subjects who cannot consider a large number of relationships simultaneously but who can decide which two of three stimuli are most similar and which two are least similar. Children as young as 5 years of age can make such judgments. Seitz (1971) demonstrated the usefulness of this type of data, using metric scaling techniques suggested by Torgerson (1958). Metric techniques, however, require repeated measurement on each triad since distance is estimated from frequency of choice. For more than five stimuli, many trials for a single subject or the data from several subjects must be combined to obtain the necessary measures. There is a major advantage in not combining the data from more than one subject since one is usually interested in a description of the perceived similarity among stimuli for individuals. The nonmetric procedures used in the present program can be employed without repeated measurements on each triad and can be applied to individual subjects. Thus, for six stimuli, only 20 trials are necessary; for nine stimuli, 84 trials are necessary.

Input. Input consists of the data gathered from all possible triads, given N stimuli. The data from each triadic stimulus array forms a data input unit consisting of the designation numbers of the two stimuli judged as most similar and the two stimuli judged as least similar. The data input units can be read into the program in any order as long as the data within each unit follows the sequence: most similar pair, least similar pair.

The Computer Science Center at The University of Maryland provided computer time for the development of this program. Reprints are available from the authors: Psychology Department, University of Maryland Baltimore County, 5401 Wilkens Avenue, Baltimore, Maryland 21228.

Output. A triangular matrix representing the subject's judgments is printed for each subject as well as Kendall's (1955) d (the total number of implied circular triads). A rank order of similarity is assigned to all the pairs; if a pair is not compared, tied scores result. These rank orders are sorted in the pair order (1,2), (1,3), ..., (1, N), ..., ($N-1$, N) which is the appropriate input for PARSCAL. These rank orders can be used directly as input for nonmetric scaling programs.

User options. The user can choose to have the similarity ranks punched for input into an appropriate scaling program. Only control cards indicating the appropriate options in the scaling program need be added.

Limitations. The program is presently limited to nine or fewer stimuli, but this is easily increased by changing dimension statements and format cards. The appropriate change points are marked by comment cards. Forty cases with six stimuli take about 7 sec on a UNIVAC 1108, nine stimuli take about 40 sec.

Computer and language. The program is written in FORTRAN IV and has been used on a UNIVAC 1108. Any system which accepts FORTRAN IV will probably accept this program.

Availability. The documented program is available from the authors at: Department of Psychology, University of Maryland Baltimore County, 5401 Wilkens Avenue, Baltimore, Maryland 21228.

REFERENCES

- COOMBS, C. H. *A theory of data*. New York: Wiley, 1964.
JOHNSON, R. M. Pairwise nonmetric multidimensional scaling. *Psychometrika*, 1973, **38**, 11-18.
KENDALL, M. G. *Rank correlation methods*. New York: Hafner Publishing Company, 1955.
KRUSKAL, J. B. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 1964, **29**, 1-27.
SEITZ, V. R. Multidimensional scaling of dimensional preferences: A methodological study. *Child Development*, 1971, **42**, 1701-1720.
SHEPARD, R. N. The analysis of proximities: Multidimensional scaling with an unknown distance function, I. *Psychometrika*, 1962, **27**, 125-140. (a)
SHEPARD, R. N. The analysis of proximities: Multidimensional scaling with an unknown distance function, II. *Psychometrika*, 1962, **27**, 219-246. (b)
SHEPARD, R. N. Metric structures in ordinal data. *Journal of Mathematical Psychology*, 1966, **3**, 287-315.
TORGERSON, W. S. *Theory and methods of scaling*. New York: Wiley, 1958.
YOUNG, F. W., & TORGERSON, W. S. TORSCA: A FORTRAN IV program for Shepard-Kruskal multidimensional scaling analysis. *Behavioral Science*, 1967, **12**, 498.

(Received for publication April 1, 1975;
revision received April 14, 1975.)