PROGRAM ABSTRACTS/ALGORITHMS

Matching factor structures using orthogonal rotation to congruence

JOHN R. VOKEY

University of Lethbridge, Lethbridge, Alberta, Canada

Of the analytic methods of transformation applied to factor-structure matrices derived from principal components analysis, *varimax* is the most widely used and, among the more common orthogonal rotation methods, generally the most satisfactory. As with these other commonly used methods, such as *quartimax* and *equamax*, the *varimax* criterion is based entirely on the mathematical consequences of an assumed *simple structure* underlying the input correlation matrix—in the case of *varimax*, the maximization of variance associated, *ceteris paribus*, with a bimodal distribution of loadings on each factor (Cliff, 1987). There is no provision in these methods for the rotation to be shaped further or otherwise limited by knowledge of factor structure arising *a priori* from the input variables and theory or from prior solutions with the same variables.

There are, however, other methods. Cliff (1966) describes a procedure for rotating principal components to maximize the match to a target structure or solution. This target may be some earlier structure of which the current solution is an attempted replication, or a factor structure founded on theoretical expectations in which the component loadings are, for example, ones and zeros. A variant of the method also may be used to match factors, so as, for example, to help determine the number of reliable components, by rotating both factor structures as closely as possible toward one another and then correlating the corresponding components. As with *varimax*, this rotation to congruence is orthogonal, maintaining the mathematical simplicity of such solutions, and it is based on the maximization of a simple criterial property—in this case, a similarity measure calculated as the sum of the products of the elements of the corresponding columns or “factors” of the two component solutions. Unlike *varimax*, however, rotation to congruence is not often used. As Cliff (1987) notes, a major reason may be that, despite its conceptual simplicity and apparent utility, rotation to congruence is not an option available in the more common statistical computing packages to which most users

have access, in contrast to the more usual techniques of *varimax*, *quartimax*, and the like.

Rotation to congruence can easily be programmed, however, if matrix-algebraic procedures are available, as with the SAS¹ MATRIX PROCedure or the EIGRS subroutine in the matrix-algebra package of IMSL² available on many mainframe computer systems. Following the exposition of Cliff (1987, pp. 328-330), with F as the matrix of component loadings to be rotated and A as the target matrix of component loadings on the same variables, the factor-structure matrix, $\hat{A}$, resulting from the rotation of F to A that maximizes the similarity between them, is given by $\hat{A} = FW_FW'_A$, in which the columns of W_F and W_A are the eigenvectors of $F'AA'F$ and $A'FF'A$, respectively. The matrix products $M_F = FW_F$ and $M_A = AW_A$ may be used to rotate the structure matrices so that the corresponding components are matched as closely as possible. Although the required matrix multiplications are simple enough, not everyone has access to the matrix-algebraic subroutine facilities, such as those provided by SAS or IMSL, to compute the eigenvectors required to use orthogonal rotation to congruence. For this reason and to facilitate the use of this method of orthogonal transformation, the program CLIFF was developed.

The Program

CLIFF is written in ZBASIC,³ the compiled BASIC language available for Macintosh, MS-DOS, Apple II, and Z80 microcomputers. With appropriate conversion of the structure and procedure facilities of ZBASIC, the source code should be easy to translate into most other dialects of the BASIC programming language. The bulk of the program is concerned with the calculation of eigenvalues and eigenvectors, which is handled by an internal procedure called EIGEN. EIGEN is based on the FORTRAN subroutine GIVHO developed by Ortega (1967), which, in turn, is based on the Givens-Householder method for computing the eigenvalues and eigenvectors of square, symmetric, real-numbered matrices. EIGEN takes a real-numbered $N \times N$, symmetric matrix, and the integer values N , NEV , and $NVEC$ ($NVEC \leq NEV \leq N$) as input parameters, diagonalizes the matrix, and returns the first NEV largest eigenvalues and the first $NVEC$ eigenvectors. As dimensioned within the program CLIFF, $N = 10$, which limits the number of factors in each structure matrix to a maximum of 10.

The remainder of the program is concerned with reading the structure matrices, stored as two separate ASCII text files on disk, and performing the matrix multiplications discussed earlier. For the latter, the program fol-

This work was supported by Operating Grant A2609 from the Natural Sciences and Engineering Research Council of Canada to the author. The assistance provided by Frank Schaffer, Seamus O'Shea, and an anonymous reviewer is greatly appreciated. Correspondence may be addressed to John R. Vokey, Department of Psychology, University of Lethbridge, Lethbridge, Alberta T1K 3M4, Canada.

lows Cliff (1966), deriving W_F from W_A and $A'F$ to reduce the processing time associated with having to execute EIGEN more than once. As currently dimensioned, the structure matrices may contain up to 20 variables. The number of factors in F may be greater than or equal to, but may not be less than that of A .⁴ All calculations are performed in double precision, but storage of intermediate results and program output are in single precision.⁵ Program output, which may be directed to the screen or a printer, consists of the original structure matrices (A and F), the orthogonal rotation of F that best matches A ($\hat{A}$), the rotations of F and A for which the corresponding columns are as similar as possible (M_F and M_A), and the correlations between these matched columns.

Availability

A printed listing of the source code, examples of structure matrices and corresponding program output, and instructions on using the program may be obtained by writing to the author. Both source-code and executable object-code files are available on diskette for Macintosh, Apple II, and MS-DOS microcomputers, and they may be obtained by writing to the author and enclosing a diskette appropriate for the operating system desired (a 3.5-in. diskette for the Macintosh version, either a 3.5-in. or a 5.25-in. diskette for the Apple II version, and a 5.25-in. diskette for the MS-DOS version). The user interface,

and the program input and output, are identical for all versions of the program.

REFERENCES

- CLIFF, N. (1966). Orthogonal rotation to congruence. *Psychometrika*, 31, 33-42.
- CLIFF, N. (1987). *Analyzing multivariate data*. San Diego, CA: Harcourt Brace Jovanovich.
- ORTEGA, J. (1967). The Givens-Householder method for symmetric matrices. In H. S. Wilf (Ed.), *Mathematical methods for digital computers* (Vol. 2: pp. 94-115). New York: Wiley.

NOTES

1. SAS Institute Inc., Box 8000 SAS Circle, Cary, NC 27511-8000.
2. IMSL, 2500 Parkwest Tower One, Houston, TX 77042-3020.
3. ZEDCOR, Inc., 4500 East Speedway Blvd., Suite 22, Tucson, AZ 85712-5305.
4. Following Cliff's (1966) original exposition of the method, the program allows the source matrix to have more columns (factors) than the target matrix. As Cliff (1987) has noted more recently, however, "experience has shown that both should be the same order" (p. 329).
5. Precision is freely adjustable in ZBASIC; for all versions of the executable object files of the program available from the author, eight digits and six digits of precision were used for double and single precision, respectively.

(Manuscript received September 27, 1988;
revision accepted for publication December 14, 1988.)