Transformation of variables to linear combinations

KEVIN E. O'GRADY

University of New Mexico, Albuquerque, New Mexico 87131

There exist many research situations in which linear combinations of an original set of measures are desirable. For instance, the calculation of subscale scores from items on a test, the selection (a priori or a posteriori) of a valid combination of a larger set of variables, or the contrasts among a set of scores may all be considered cases for which a solution lies in the application of a suitable transformation matrix to the observations of each subject. Indeed, the research design for which a linear combination of the original variable(s) would prove most helpful is the frequently used one in which each subject is repeatedly sampled on one or more criterion variables, what is often referred to as a "within-subjects" or "repeated-battery" design. In this case, the application of an appropriate contrast matrix to the outcome measure(s) produces linear combinations of the original variable(s) that can be used in profile analysis or growth-curve analysis (for example, see Harris, 1975; Timm, 1975).

Description. The programs consist of a brief main routine and several subroutines. The MAIN routine simply controls program flow. POLYN reads the metric of the polynomial contrast matrix, if the user has indicated the metric is unequally spaced, and then calculates the polynomial contrast matrix. USERD reads a user-defined contrast matrix. ORTHO performs a Gram-Schmidt orthonormalization on the rows of polynomial contrast matrix, or, if requested by the user, on the rows of the user-defined contrast matrix. INPUT reads the input data, subject by subject, and applies the contrast matrix to it. The transformed data are temporarily stored in binary form on a peripheral device. RORDER reorders the input data prior to the application of the contrast matrix, if requested by the user. OUTPUT prints the contrast matrix applied to the data, reads and prints, and punches and/or writes to an alternate medium the transformed data, as requested by the user.

Input. Input to the program consists of (1) a title card, containing any 80 alphanumeric characters, (2) a parameter card, indicating the number of subjects, number of variables, number of input and output format cards, type of transformation matrix to be read or generated, and various other program options, (3) variable format card(s) describing the input data, (4) variable

format card(s) indicating how the transformed data are to be written, (5) the input data. The input data may reside on some medium other than cards that follow immediately in the input card stream.

Output. The printed output includes the job title, program options selected, and the transformation matrix applied to the input data. It is also possible to have the transformed data listed on the printed output. In addition, the transformed data can be written onto an alternate output device (e.g., tape, disk), punched onto cards, or both.

Technical Information. The program is written in 1966 ANSI standard FORTRAN IV and should be compatible with any mainframe or minicomputer FORTRAN compiler with few if any modifications. All computations are in single-precision arithmetic, and core storage requirements on an IBM 3032 are approximately 70 KB. This includes 30 KB for program storage and 40 KB for array storage. Array storage consists entirely of a singly subscripted array stored in COMMON. This array is subdivided at execution time according to the various program options chosen, through the use of "integer pointers" that allocate various portions of the array. Thus, the program is quite flexible in terms of the variety of programs it can handle. The formula detailing the limitations in core storage is complicated and, thus, is not presented here. As a rough guideline, the number of observations available on each subject must not exceed 100. This limitation can be altered easily by changing three statements in the MAIN routine.

Moreover, one peripheral storage device is used as a binary scratch unit, and one peripheral device can be allocated for permanent storage of the output data set. The units numbers for these devices, as well as the card reader, printer, punch, and an alternate device that might contain the input data set, are assignable within the BLOCK DATA subprogram. Currently, these devices are assigned the numbers 1, 4, 5, 6, 7, and 0, respectively.

Availability. Reprints of this paper and an annotated source listing can be obtained at no charge from Kevin E. O'Grady, Department of Psychology, University of New Mexico, Albuquerque, New Mexico 87131.

REFERENCES

- HARRIS, R. J. *A primer of multivariate statistics*. New York: Academic Press, 1975.
TIMM, N. H. *Multivariate analysis with applications in education and psychology*. Monterey, Calif: Brooke/Cole, 1975.

(Accepted for publication November 16, 1982.)