SOME MATHEMATICAL NOTES ON THREE-MODE FACTOR ANALYSIS*

LEDYARD R. TUCKER
UNIVERSITY OF ILLINOIS

The model for three-mode factor analysis is discussed in terms of newer applications of mathematical processes including a type of matrix process termed the Kronecker product and the definition of combination variables. Three methods of analysis to a type of extension of principal components analysis are discussed. Methods II and III are applicable to analysis of data collected for a large sample of individuals. An extension of the model is described in which allowance is made for unique variance for each combination variable when the data are collected for a large sample of individuals.

Extension of the two-mode factor analytic model to three or more modes of data classification has been suggested by Tucker. Initial discussions of this development appear in the monographs: Problems in Measuring Change [8] and Contributions to Mathematical Psychology [9]. The latter of these two monographs gives the basic mathematical structure of the proposed model. A further discussion of the mathematical structure was given by Levin in his PhD dissertation Three-mode factor analysis [4]. Results of experimental trials of the method were reviewed by Tucker in a paper read at the 1964 Invitational Conference on Testing Problems [10]. Since the Tucker and Levin descriptions of the mathematical structure of the model and analysis procedures, there have been several mathematical developments which add power and clarity to the structure of the model. The structure of the three-mode factor analytic model is discussed here in terms of the newer mathematical statements. A further refinement to be considered involves allowances for a type of unique variance related to errors of measurement. A fictitious body of data is used to illustrate several points.

Remarks on Notation

In the development of the three-mode factor analysis model it has been found quite useful to adopt several rather unique features of notation. Some of these notational items are at variance with common mathematical usage but have been found helpful in consideration of some relatively complex relations. Much of standard summational and matrix notation has been

* This research was jointly supported by the University of Illinois and the Office of Naval Research under contract Nonr 1834 (39). The author is most appreciative of the careful reading of the manuscript by Dr. Douglas McKie and Dr. Maurice Tatsuoka.

279
A first item is the use of the word \textit{mode}. Tucker introduced this term to denote "a set of indices by which data might be classified" ([9], p. 112). For example, the scores of a sample of individuals on a battery of tests could be classified by the individuals in the sample and cross-classified by the tests in the battery. The individuals in the sample would be the elements of one set of indices by which the scores are classified; thus, the sample of individuals would constitute one mode of the data. A second mode of this data would be the battery of tests. The test scores could be arranged in a rectangular table with rows for individuals and columns for tests. Such an arrangement will be termed, in the present context, a two-mode matrix. If the battery of tests were administered to the sample of individuals on several occasions, the set of occasions would be considered as a third mode. The data, now, could be arranged in a rectangular prism or box with horizontal strata of cells for individuals, vertical strata parallel to the end planes for tests, and vertical strata parallel to the front plane for occasions. Such an arrangement will be termed a three-mode matrix. In general, an \(n\)-mode matrix would involve cross-classification of the data on \(n\) sets of indices, or modes. Each datum would correspond to an element of the Cartesian product of the sets of classification indices or modes.

Each mode will be identified by a lower-case letter, for example, the letter \(i\) may be used for the mode for individuals in a sample. It has proven convenient to use this lower case letter in several related, but distinct roles: 1) as a general identification of the mode, 2) as a subscript identifying the mode to which an element belongs, and 3) as a variable identification symbol for the elements in the mode. An example of the first usage is a statement such as "mode \(i\) is for the individuals in the sample." An example of the second usage is in the assignment of identification symbols \(1_i, 2_i, 3_i, \ldots, N_i\), to the individuals in the sample. The identification symbol for each element of a mode is composed to two parts, one part being a number termed the index value of the element, designated by \(v(i)\), and the other part being the identification subscript for mode. It will be noted that the elements for each mode constitute an ordered set. The index value will be utilized in any calculations for identification of elements. The series of index values for the elements in a mode shall consist of the integers from 1 to the number of elements in the mode. The number of elements in a mode will be designated by the capital letter \(N\) with the subscript identifying the mode, that is, by \(N_i\), where \(m\) is used in the present context as a generalized mode identification; thus, \(N_i\) is the number of elements in mode \(i\). In the third role, the letter is used as a general, unspecified identification symbol which may be particularized to the identification symbol of each of the elements in turn. For example, \(x_{i,j,k}\) will be used as the generalized entry in the three-mode matrix.