A NOTE ON INVARIANCE IN THREE-MODE FACTOR ANALYSIS

Bruce Bloxom

Educational Testing Service

Previous results of the application of Lawley's selection theorem to the common factor analysis model are extended to a revision of Tucker's three-mode principal components model. If the regression of the three-mode manifest variates on variates used to select subpopulations is both linear and homoscedastic, the two factor pattern matrices, the core matrix, and the residual variance-covariance matrix in the three-mode model can all be assumed to be invariant across subpopulations. The implication of this finding for simple structure is discussed.

A question of major importance for factor analysis is centered on the problem of factorial invariance. Meredith [1964] reviewed the problem with respect to the common factor analysis model. He found that under certain conditions the factor pattern matrix and the residual variance-covariance matrix are invariant across subpopulations.

Tucker [1963] proposed a principal components model for data where measurement is taken in two "modes" over a set of subjects. The model includes (a) a matrix of parameters for each of the two measurement modes, (b) a matrix of parameters for the persons mode, (c) a three-mode core matrix which relates the measurement modes to the persons mode, and (d) a three-mode error matrix. If that model is written in terms of random variables instead of a finite matrix of person parameters, the same question can be asked about the model as is asked about the common factor analysis model: Which of the matrices can be assumed to be invariant across subpopulations? The present paper extends Meredith's results with the common factor analysis model to a three-mode random factor analysis model and discusses the resulting implications for simple structure.

Three-Mode Factor Invariance

Consider a set of $j$ measures (e.g., a set of ratings) taken on a set of $k$ objects (e.g., a set of stimuli with various physical properties) for a population of persons. This defines a $j \times k$-dimensional random vector,

$$y = \{y_{11}, \cdots, y_{1k}, y_{21}, \cdots, y_{2k}, \cdots, y_{i1}, \cdots, y_{ik}\}.$$  

Suppose that $y$ is related to some underlying $\ell$-dimensional vector of factor scores, $z$, so that
where $\beta$ is a factor pattern matrix and $u$ is a vector of residual variates. Assuming $z$ and $u$ to be independent, and $E(u'u)$ to be diagonal, we have the common factor analysis model

$$E(y'y) = V = \beta' \phi \beta + \Delta,$$

where $V$, $\phi$, and $\Delta$ are expected variance-covariance matrices of $y$, $z$, and $u$ respectively.

Meredith [1964] has shown that if the assumptions of Lawley’s [1943] selection theorem can be met, $\beta$ and $\Delta$ in (2) are invariant across subpopulations. These assumptions are (a) linearity and (b) homoscedasticity of regression of the manifest variates, $y$, on variates used to select subpopulations. Then

$$V_* = \beta'\phi_*\beta + \Delta,$$

where $V_*$ is a subpopulation’s manifest variance-covariance matrix and $\phi_*$ is the variance-covariance matrix for the subpopulation’s factor scores. $V_*$ and $\phi_*$ are assumed to be of the same rank as $V$ and $\phi$ respectively.

Tucker [1963, 1966] proposed a three-mode principal components analysis model which would suggest expressing (2) as

$$y = z G (B' \times C') + u,$$

where $B'$ is a matrix of factor loadings for the $j$ measures, rank $(B) = m$, and $m < j$. $C'$ is a matrix of factor loadings for the $k$ objects, rank $(C) = n$, and $n < k$. $(B' \times C')$ is the $(m \times n)$ by $(j \times k)$ Kronecker product [Bellman, 1960] of $B'$ and $C'$. $G$ is a matrix whose indices show the relative extent to which a given row vector of $B'$ and a given row vector of $C'$ determine $y$ when combined with a given dimension in $z$.

This model differs from Tucker’s model in that $y$, $z$, and $u$ are random variables instead of matrices of parameters for a finite number of individuals. Now assuming, as in common factor analysis, that $E(u'z) = 0$ and that $E(u'u)$ is diagonal, we can substitute (5) in (3) yielding, for the population,

$$V = (B \times C)G'\phi G (B' \times C') + \Delta.$$

Let subpopulations be selected on a set of variates, $\alpha$, and assume the regression of $y$ on $\alpha$ is both linear and homoscedastic. Then from (6) and (4),

$$V_* = (B \times C)G'\phi_* G (B' \times C') + \Delta.$$

Providing the assumptions just stated are met, $B$, $C$, $G_*$, and $\Delta$ are invariant across subpopulations.