STRUCTURAL ANALYSIS OF COVARIANCE AND CORRELATION MATRICES

KARL G. JÖRESKOG

UNIVERSITY OF UPPSALA

A general approach to the analysis of covariance structures is considered, in which the variances and covariances or correlations of the observed variables are directly expressed in terms of the parameters of interest. The statistical problems of identification, estimation and testing of such covariance or correlation structures are discussed.

Several different types of covariance structures are considered as special cases of the general model. These include models for sets of congeneric tests, models for confirmatory and exploratory factor analysis, models for estimation of variance and covariance components, regression models with measurement errors, path analysis models, simplex and circumplex models. Many of the different types of covariance structures are illustrated by means of real data.

Key words: covariance structure analysis, factor analysis, variance components, path analysis, simplex, circumplex.

1. Introduction

The search for structure in correlated psychological variables has been one of the main objectives in psychometrics for several decades. Traditionally this search was done by using factor analysis to detect and assess latent sources of variation and covariation in observed measurements. Seldom do these measurements represent pure psychological traits or functions. Rather, as Thurstone [1947] assumed in his multiple factor model, each measure depends on a limited number of traits or functions and one tries to identify, and ultimately estimate, the components of the observed measurements associated with different traits or functions.

In factor analysis the correlation matrix is subjected to a suitable method for estimation of the factor space, the solution rotated to obtain projections of the test vectors on certain reference vectors, called factors, and, by examining the contents of the tests which have large projections on a particular reference vector, a trait or function is inferred to be common to these psychological tests. The trait or function, treated as an explanatory variable is then named and considered to be a source of one of the components of covariation or correlation in the tests analyzed. Individual differences in this component can then be estimated as so called factor scores.

That exploratory factor analysis may be quite useful in the early stages of experimentation or test development is widely recognized. Thurstone's [1938] primary mental abilities, French's [1951] factors in aptitude and achievement tests and Guilford's [1956] structure of intelligence are good examples of this. The results of an exploratory analysis may have heuristic and suggestive value [Anderson, 1963] and may generate hypotheses which are capable of more objective testing by other multivariate methods.

As more knowledge is gained about the nature of psychological measurements, however, exploratory factor analysis may not be a useful tool and may even become a

hindrance. Often there is structure in the data which can be postulated a priori and this structure may not be consistent with the factor analysis model at all. Such structure may arise because of a specified theory or hypothesis, a given classificatory design for items or subtests according to objective features of content and format, known experimental conditions or because of results from previous studies based on extensive data. Sometimes the observed variables are ordered through time, as in longitudinal studies, or according to linear or circular patterns, as in Guttman's [1954] simplex and circumplex models, or according to a given causal scheme, as in path analysis or structural equation models or the variables are classified into the two categories dependent and independent variables as in prediction studies. The methods of exploratory factor analysis cannot take such a given structure into account and if applied to data having such a structure, it will usually give very misleading results.

In this paper I shall describe a new general method for structural analysis of covariance and correlation matrices. The new method is in the spirit of previous writers on analysis of covariance structures, e.g. Bock [1960], Bock and Bargmann [1966], Anderson [1969], Jöreskog [1970a, 1973, 1974], Mukherjee [1970], Browne [1974, 1977] and McDonald [1974, 1975] but is based on a more general model than all previous models in that it can handle *any* linear or non-linear covariance structure. Most of the previous models for analysis of covariance structures are based on models in which Σ, the covariance matrix of the observed variables, is generated as a sum of products of matrices whose elements are either fixed a priori or free parameters to be estimated. However, often each element of Σ is a simple function of just a few of the parameters and Σ can therefore be computed directly and much more rapidly by using the specific nature of the covariance structure. The identification, estimation and testing of such a covariance structure is discussed as well as the testing of structural hypotheses within a model. Several examples are given of various models useful in the behavioral sciences.

Most studies are to some extent both exploratory and confirmatory since they involve some variables of known and other variables of unknown composition. The former should be chosen with great care in order that as much information as possible about the latter may be extracted. It is highly desirable that a hypothesis which has been suggested by mainly exploratory procedures should subsequently be confirmed, or disproved, by obtaining new data and subjecting these to more rigorous statistical techniques. Although the new method is most useful in confirmatory studies, it can also be used to do exploratory analysis by means of a nested sequence of confirmatory analyses.

2. The General Model

General Covariance and Correlation Structures

Any covariance structure may be defined by specifying that the population variances and covariances of the observed variables are certain functions of parameters $\theta_1, \theta_2, \cdots, \theta_t$ to be estimated from data: $\sigma_{ij} = \sigma_{ij}(\theta)$, or in matrix form $\Sigma = \Sigma(\theta)$. It is assumed that the functions $\sigma_{ij}(\theta)$ are continuous and have continuous first derivatives and that Σ is positive definite at every point θ of the admissible parameter space. The distribution of the observed variables is assumed to be multivariate with an unconstrained mean vector μ and covariance matrix $\Sigma(\theta)$ and is assumed to be sufficiently well described by the moments of first and second order, so that additional information about θ contained in moments of higher order may be ignored. In particular this will hold if the distribution is multivariate normal.

A correlation structure is defined by specifying that the population correlations ρ_{ij} of the observed variables are functions $\rho_{ij} = \rho_{ij}(\theta)$ of θ. Such a correlation structure is treated

as a covariance structure by specifying that

$$\text{(1)} \qquad \Sigma = D_\sigma P(\theta) D_\sigma,$$

where D_σ is a diagonal matrix of population standard deviations $\sigma_1, \sigma_2, \cdots, \sigma_p$ of the observed variables, which are regarded as free parameters, and $P(\theta)$ is the correlation matrix. The covariance structure (1) has parameters $\sigma_1, \sigma_2, \cdots, \sigma_p, \theta_1, \theta_2, \cdots, \theta_t$. The standard deviations $\sigma_1, \sigma_2, \cdots, \sigma_p$ as well as θ must be estimated from data and the estimate of σ_i does not necessarily equal the corresponding standard deviation in the sample.

Identification

Before an attempt is made to estimate the parameters θ, the identification problem must be resolved. The identification problem is essentially whether or not θ is uniquely determined by Σ. Every θ in the admissible parameter space generates a Σ but two or more θ's may possibly generate the same Σ. The whole model is said to be identified if for any two vectors θ_1 and θ_2 in a region of the parameter space, locally or globally, $\theta_1 \neq \theta_2$ implies that $\Sigma(\theta_1) \neq \Sigma(\theta_2)$, i.e. if Σ is generated by one and only one θ. This means that all parameters are identified. However, even if the whole model is not identified some parameters can still be identified. Consider the set of all parameter vectors θ generating the same Σ. If a parameter θ_i has the same value in all such vectors, this parameter is identified. For parameters which are identified the methods to be described will yield consistent estimators. If a model is not completely identified, appropriate restrictions may be imposed on θ to make it so, and the choice of restrictions may affect the interpretation of the results of an estimated model.

Identifiability depends on the choice of model. To examine the identification problem for a particular model consider the equations

$$\text{(2)} \qquad \sigma_{ij} = \sigma_{ij}(\theta), \qquad i \leq j.$$

There are $\frac{1}{2} p(p + 1)$ equations in t unknown parameters θ. Hence a necessary condition for identification of all parameters is that

$$\text{(3)} \qquad t \leq \tfrac{1}{2} p(p + 1).$$

If a parameter θ can be determined from Σ by solving the equations (2) or a subset of them, this parameter is identified; otherwise it is not. Often some parameters can be determined from Σ in several ways, i.e., by using different sets of equations. This gives rise to overidentifying conditions on Σ which must hold if the model is true. Since the equations (2) are often non-linear, the solution of the equations is often complicated and tedious and explicit solutions for all θ's seldom exist. Examples on how the identification problem is resolved in particular cases are given in Sections 3–10 of the paper.

There are various ways in which the computer program may be used to check the identification status of the model. If the GLS or ML methods are used for estimation (see the next subsection) the information matrix may be obtained and checked for positive definiteness. If the model is identified then the information matrix is almost certainly positive definite. If the information matrix is singular, the model is not identified and the rank of the information matrix may indicate which parameters are not identified. Another procedure which may also be used when other methods of estimation are used is the following. Choose a set of reasonable values for the parameters and compute Σ. Then run the program with this Σ as input matrix and estimate θ. If this results in the same estimated values as were used to generate Σ, then it is most likely that the model is identified. Otherwise, those parameters which gave a different value are probably not identified.

Estimation

The population is characterized by the mean vector μ, which is unconstrained, and the covariance matrix Σ which is a function of θ. In practice θ is unknown and must be estimated from a sample of N independent observations on the random vector x of order p. Let $S = (s_{ij})$ be the usual sample covariance matrix of order $p \times p$, based on $n = N - 1$ degrees of freedom. The information provided by S may also be represented by a correlation matrix $R = (r_{ij})$ and a set of standard deviations $s_1, s_2, \cdots, s_p$, where $s_i = (s_{ii})^{1/2}$ and $r_{ij} = s_{ij}/s_i s_j$. In many applications both the origin and the unit in the scales of measurement are arbitrary or irrelevant and then only the correlation matrix may be of any interest. In such cases one takes S to be the correlation matrix R in what follows.

Since the mean vector is unconstrained, and higher moments are ignored, the estimation problem is how to fit a matrix Σ of the form $\Sigma(\theta)$ to the observed covariance matrix S. Three different methods of fitting Σ to S will be considered, namely the *unweighted least squares* (ULS) method, which minimizes

$$(4) \qquad U = \tfrac{1}{2} \operatorname{tr} (S - \Sigma)^2,$$

the *generalized least squares* (GLS) method, which minimizes

$$(5) \qquad G = \tfrac{1}{2} \operatorname{tr} (I - S^{-1}\Sigma)^2,$$

and the *maximum likelihood* (ML) method, which minimizes

$$(6) \qquad M = \operatorname{tr} (\Sigma^{-1}S) - \log |\Sigma^{-1}S| - p.$$

Each function is to be minimized with respect to θ.

All three functions U, G and M may be minimized by basically the same algorithm. The notation $F = F(S, \Sigma)$ will be used for any one of the three functions. The GLS and ML methods are scale-free in the sense that $F(S, \Sigma) = F(DSD, D\Sigma D)$, for any diagonal matrix of positive scale factors; ULS does not have this property. With ULS, an analysis of S and of DSD yield results which may not be properly related. When x has a multivariate normal distribution both GLS and ML yield estimates that are efficient in large samples. Both GLS and ML require a positive definite covariance matrix S or correlation matrix R; ULS will work even on a matrix which is non-gramian.

Under the assumption that x has a multinormal distribution or that S has a Wishart distribution, M in (6) is a transform of the log-likelihood function for the sample, hence its association to the maximum likelihood method. Jöreskog and Goldberger [1972] derived the expression for G from Aitken's [1934–35] principle of generalized least squares using estimated asymptotic variances and covariances of the elements of S under multinormality of x. Browne [1974] justified GLS under the slightly more general assumption that the elements of S have an asymptotic normal distribution. Since the variances and covariances in S are generally correlated and have unequal variances, it would seem that ULS uses the wrong metric in measuring deviations between S and Σ. Nevertheless, ULS produces consistent estimators under more general assumptions than those which have been used to justify ML and GLS.

The derivatives of F are

$$(7) \qquad \frac{\partial F}{\partial \theta_i} = \operatorname{tr} \left[A(\Sigma - S)A \, \frac{\partial \Sigma}{\partial \theta_i} \right],$$

where $A = I$ in ULS, $A = S^{-1}$ in GLS and $A = \Sigma^{-1}$ in ML. Assuming that S converges in probability to Σ and ignoring terms of order $\Sigma - S$, the second derivatives are approximately

$$(8) \qquad \frac{\partial^2 F}{\partial \theta_i \, \partial \theta_j} = \operatorname{tr} \left[\frac{A\partial \Sigma}{\partial \theta_i} \, \frac{A\partial \Sigma}{\partial \theta_j} \right].$$

Note that both (7) and (8) may be computed from knowledge about only the first derivatives of the covariance structure functions $\Sigma(\theta)$. In ML and GLS, (8) yields the elements of the information matrix which is positive definite at every point θ of the admissible parameter space, if θ is identified.

The function $F(\theta)$ may be minimized numerically by Fisher's scoring method [see e.g. Rao, 1973, Section 5g] or the method of Fletcher and Powell [1963]; see also Gruvaeus and Jöreskog [Note 1].

The minimization starts at an arbitrary starting point $\theta^{(1)}$ and generates successively new points $\theta^{(2)}$, $\theta^{(3)}$, $\cdots$, such that $F(\theta^{(s+1)}) < F(\theta^{(s)})$ until convergence is obtained.

Let $\mathbf{g}^{(s)}$ be the gradient vector $\partial F / \partial \theta$ at $\theta = \theta^{(s)}$ and let $\mathbf{E}^{(s)}$ be the information matrix given by (8) and evaluated at $\theta = \theta^{(s)}$. Then Fisher's scoring method computes a correction vector $\boldsymbol{\delta}^{(s)}$ by solving the equation system

$$(9) \qquad \mathbf{E}^{(s)}\boldsymbol{\delta}^{(s)} = \mathbf{g}^{(s)}$$

and then computes the new point as

$$(10) \qquad \theta^{(s+1)} = \theta^{(s)} - \boldsymbol{\delta}^{(s)}.$$

This requires the computation of $\mathbf{E}^{(s)}$ and the solution of (9) in each iteration, and this is often quite time consuming. An alternative is to use the method of Fletcher and Powell, which evaluates only the inverse of $\mathbf{E}^{(1)}$ and in subsequent iterations $\mathbf{E}^{-1}$ is improved, using information built up about the function, so that ultimately $\mathbf{E}^{-1}$ converges to an approximation of the inverse of $\partial^2 F / \partial \theta\, \partial \theta'$ at the minimum.

In GLS and ML, $(2/N)$ times the inverse of the information matrix $\mathbf{E}$, given by (8) and evaluated at the minimum of F, provides an estimate of the asymptotic covariance matrix of the estimators $\hat{\theta}$ of θ. The square root of the diagonal elements of $(2/N)\mathbf{E}^{-1}$ are large-sample estimates of the standard errors of the $\hat{\theta}$'s.

Unfortunately no statistical theory is available for computing standard errors for ULS estimators. Such standard errors may be obtained by jackknifing but this requires extensive computation.

Assessment of Fit

When the number of independent parameters in θ is less than the total number of variances and covariances in Σ, i.e., when $t < \frac{1}{2}p(p + 1)$, the model imposes conditions on Σ which must hold if the model is true. In GLS and ML, the validity of these conditions, i.e., the validity of the model, may be tested by a likelihood ratio test. The logarithm of the likelihood ratio is simply $(N/2)$ times the minimum value of the function F. Under the model, this is distributed, in large samples, as a χ^2 distribution with degrees of freedom equal to

$$(11) \qquad d = \frac{1}{2}p(p + 1) - t$$

Tests of Structural Hypotheses

Once the validity of the model has been reasonably well established, various structural hypotheses within the model may be tested. One can test hypotheses of the forms

(i) that certain θ's are fixed equal to assigned values and/or
(ii) that certain θ's are equal in groups.

Each of these two types of hypotheses leads to a covariance structure $\Sigma(v)$ where v is a subset of $u < t$ elements of θ. Let F_v be the minimum of F under the structural hypothesis and let F_θ be the minimum of F under the general model. Then $(N/2)(F_v - F_\theta)$ is approximately distributed as χ^2 with $t - u$ degrees of freedom.

The Use of χ^2 in Exploratory Studies

The values of χ^2 should be interpreted very cautiously because of the sensitivity of χ^2 to various model assumptions such as linearity, additivity, multinormality, etc., but also for other reasons. In most empirical work many of the models considered may not be very realistic. If a sufficiently large sample were obtained, the test statistic would, no doubt, indicate that any such model is statistically untenable. The model should rather be that $\Sigma(\theta)$ represents a reasonable approximation to the population covariance matrix. From this point of view the statistical problem is not one of testing a given hypothesis (which a priori may be considered false) but rather one of fitting various models with different numbers of parameters and to decide when to stop fitting. In other words, the problem is to extract as much information as possible out of a sample of given size without going so far that the result is affected to a large extent by "noise". It is reasonable and likely that more information can be extracted from a large sample than from a small one. In such a problem it is the difference between χ^2 values that matters rather than the χ^2 values themselves. In an exploratory study, if a value of χ^2 is obtained which is large compared to the number of degrees of freedom, the fit may be examined by an inspection of the residuals, i.e., the discrepancies between observed and reproduced values. Often the results of an analysis, an inspection of residuals or other considerations will suggest ways to relax the model somewhat by introducing more parameters. The new model usually yields a smaller χ^2. If the drop in χ^2 is large compared to the difference in degrees of freedom, this is an indication that the change made in the model represents a real improvement. If, on the other hand, the drop in χ^2 is close to the difference in number of degrees of freedom, this is an indication that the improvement in fit is obtained by "capitalizing on chance" and the added parameters may not have any real significance or meaning.

Often it is not possible, or even desirable, to specify the model completely since there may be other models which are equally plausible. In such a situation it is necessary to have a technique of analysis which will give information about which of a number of alternative models is (are) the most reasonable. Also, if there is sufficient evidence to reject a given model due to poor fit to the data, the technique should be designed to suggest which part of the model is causing the poor fit. Several examples will be given illustrating the assessment of fit of a model and strategies for model modification.

In the following sections of this paper, several examples are given of models that are useful in the behavioral sciences and some of these models are illustrated by means of real data. All analyses presented are based on the ML method and all χ^2 values have been obtained as $(N - 1)$ times the minimum value of M. Sometimes a probability level P for a χ^2 value is given. This refers to the probability of obtaining a χ^2 larger than that actually obtained, given that the hypothesized model holds.

3. Models for Sets of Congeneric Tests

Test Theory Models

Most measurements employed in the behavioral sciences contain sizeable errors of measurements and any adequate theory or model must take this fact into account. Of particular importance is the study of congeneric measurements, i.e., those measurements that are assumed to measure the same thing.

Classical test theory [Lord, & Novick, 1968] assumes that a test score x is the sum of a true score τ and an error score e, where e and τ are uncorrelated. A set of test scores x_1, $\cdots$, x_p with true scores τ_1, $\cdots$, τ_p is said to be congeneric if every pair of true scores τ_i and τ_j have unit correlation. Such a set of test scores can be represented as

$$\mathbf{x} = \mathbf{\mu} + \mathbf{\beta}\tau + \mathbf{e},$$

where $\mathbf{x}' = (x_1, \cdots, x_p)$, $\boldsymbol{\beta}' = (\beta_1, \cdots, \beta_p)$ is a vector of regression coefficients, $\mathbf{e}' = (e_1, \cdots, e_p)$ is the vector of error scores, $\boldsymbol{\mu}$ is the mean vector of $\mathbf{x}$ and $\boldsymbol{\tau}$ is a true score, for convenience scaled to zero mean and unit variance. The elements of $\mathbf{x}$, $\mathbf{e}$ and $\boldsymbol{\tau}$ are regarded as random variables for a population of examinees. Let $\theta_1^2, \cdots, \theta_p^2$ be the variances of $e_1, \cdots, e_p$, respectively, i.e., the error variances. The corresponding true score variances are $\beta_1^2, \cdots, \beta_p^2$. One important problem is that of estimating these quantities. The covariance matrix of $\mathbf{x}$ is

(12) $$\boldsymbol{\Sigma} = \boldsymbol{\beta}\boldsymbol{\beta}' + \boldsymbol{\Theta}^2,$$

where

$$\boldsymbol{\Theta} = \text{diag }(\theta_1, \cdots, \theta_p).$$

Parallel tests and tau-equivalent tests, in the sense of Lord and Novick [1968], are special cases of congeneric tests. Parallel tests have equal true score variances and equal error variances, i.e.,

$$\beta_1^2 = \cdots = \beta_p^2, \qquad \theta_1^2 = \cdots = \theta_p^2.$$

Tau-equivalent tests have equal true score variances but possibly different error variances.

Parallel and tau-equivalent tests are homogenous in the sense that all covariances between pairs of test scores are equal. For parallel tests the variances are also equal. Scores on such tests are directly comparable, i.e., they represent measurements on the same scale. For tests composed of binary items this can hold only if the tests have the same number of items and are administrated under the same time limits. Congeneric tests, on the other hand, need not satisfy such strong restrictions. They need not even be tests consisting of items but can be ratings, for example, or even measurements produced by different measuring instruments.

Recently Kristof [1971] developed a model for tests which differ only in length. This model assumes that there is a "length" parameter β_i associated with each test score x_i in such a way that the true score variance is proportional to β_i^4 and the error variance proportional to β_i^2. It can be shown that the covariance structure for this model is of the form

(13) $$\boldsymbol{\Sigma} = \mathbf{D}_\beta(\boldsymbol{\beta}\boldsymbol{\beta}' + \psi^2\mathbf{I})\mathbf{D}_\beta,$$

where $\mathbf{D}_\beta = \text{diag }(\beta_1, \beta_2, \cdots, \beta_p)$ and $\boldsymbol{\beta}' = (\beta_1, \beta_2, \cdots, \beta_p)$. This model has $p + 1$ independent parameters and is less restrictive than the parallel model but more restrictive than the congeneric model. A summary of the various test theory models and their number of parameters is given in Table 1.

As an illustration of the variable-length model consider the following covariance matrix $\mathbf{S}$ taken from Kristof [1971]:

$$\mathbf{S} = \begin{bmatrix} 54.85 & & \\ 60.21 & 99.24 & \\ 48.42 & 67.00 & 63.81 \end{bmatrix}.$$

This is based on $N = 900$ candidates who took the January 1969 administration of the Scholastic Aptitude Test (SAT). The first test, Verbal Omnibus, was administered under 30 minutes, and the second test, Reading Comprehension, under 45 minutes, these two tests having 40 and 50 items respectively. The third test is an additional section of the SAT, not normally administered.

The following maximum likelihood estimates are obtained: $\hat{\beta}_1 = 2.58$, $\hat{\beta}_2 = 3.03$, $\hat{\beta}_3 = $

TABLE 1

Various Test Theory Models

Model	Covariance Structure	No. of Parameters
Parallel	$\underset{\sim}{\Sigma} = \beta^2 \underset{\sim}{j}\underset{\sim}{j}' + \theta^2 \underset{\sim}{I}$	2
Tau-equivalent	$\underset{\sim}{\Sigma} = \beta^2 \underset{\sim}{j}\underset{\sim}{j}' + \underset{\sim}{\Theta}^2$	$p + 1$
Variable-length	$\underset{\sim}{\Sigma} = \underset{\sim}{D}_\beta (\beta\beta' + \psi^2 \underset{\sim}{I}) \underset{\sim}{D}_\beta$	$p + 1$
Congeneric	$\underset{\sim}{\Sigma} = \underset{\sim}{\beta}\underset{\sim}{\beta}' + \underset{\sim}{\Theta}^2$	$2p$

$\underset{\sim}{j}$ denotes a column vector with all elements equal to one.

2.69 and $\hat{\psi} = 1.60$. The goodness of fit test yield $\chi^2 = 4.93$ with 2 degrees of freedom. This has a probability level of 0.09.

Several Sets of Congeneric Test Scores

The previous model generalizes immediately to several sets of congeneric test scores. If there are q sets of such tests, with $m_1, m_2, \cdots, m_q$ tests respectively, we write $x' = (x'_1, x'_2, \cdots, x'_q)$ where x'_g, $g = 1, 2, \cdots, q$ is the vector of observed scores for the gth set. Associated with the vector x_g there is a true score τ_g and vectors μ_g and β_g defined as in the previous section so that

$$x_g = \mu_g + \beta_g \tau_g + e_g.$$

As before we may, without loss of generality, assume that τ_g is scaled to zero mean and unit variance. If the different true scores $\tau_1, \tau_2, \cdots, \tau_q$ are all mutually uncorrelated, then each set of tests can be analyzed separately as in the previous section. However, in most cases these true scores correlate with each other and an overall analysis of the entire set of tests must be made. Let $p = m_1 + m_2 + \cdots m_q$ be the total number of tests. Then x is of order p. Let μ be the mean vector of x, and let e be the vector of error scores. Furthermore, let

$$\tau' = (\tau_1, \tau_2, \cdots, \tau_q)$$

and let B be the matrix of order $p \times q$, partitioned as

$$B = \begin{bmatrix} \beta_1 & 0 & \cdots & 0 \\ 0 & \beta_2 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & \beta_q \end{bmatrix}$$

Then x is represented as

(14) $x = \mu + B\tau + e.$

Let Γ be the correlation matrix of τ. Then the covariance matrix Σ of x is

$$(15) \qquad\qquad \Sigma = B\Gamma B' + \Theta^2,$$

where Θ^2 is a diagonal matrix of order p containing the error variances.

Estimation and Testing of Disattenuated Correlation Coefficients

The correlation·coefficient corrected for attenuation between two tests x and y is the correlation between their true scores. If, on the basis of a sample of examinees, the disattenuated coefficient is near unity, the experimenter concludes that the two tests are measuring the same trait.

Let x_1, x_2, y_1, and y_2 be four tests with zero means and satisfying the following model

$$\begin{bmatrix} x_1 \\ x_2 \\ y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \beta_1 & 0 \\ \beta_2 & 0 \\ 0 & \beta_3 \\ 0 & \beta_4 \end{bmatrix} \begin{bmatrix} \tau_x \\ \tau_y \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{bmatrix}$$

with covariance matrix

$$\Sigma = \begin{bmatrix} \beta_1 & 0 \\ \beta_2 & 0 \\ 0 & \beta_3 \\ 0 & \beta_4 \end{bmatrix} \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \begin{bmatrix} \beta_1 & \beta_2 & 0 & 0 \\ 0 & 0 & \beta_3 & \beta_4 \end{bmatrix} + \begin{bmatrix} \theta_1^2 & 0 & 0 & 0 \\ 0 & \theta_2^2 & 0 & 0 \\ 0 & 0 & \theta_3^2 & 0 \\ 0 & 0 & 0 & \theta_4^2 \end{bmatrix}$$

In this model, x_1 and x_2 are congeneric measures of τ_x and y_1 and y_2 are congeneric measures of τ_y. The disattenuated correlation ρ is the correlation between τ_x and τ_y. One would like to estimate ρ and test the hypothesis that $\rho = 1$.

To illustrate this some data from Lord [1957] are used. The tests x_1 and x_2 are 15-item vocabulary tests administered under liberal time limits. The tests y_1 and y_2 are highly speeded 75-item vocabulary tests. The covariance matrix is given in Table 2a. The data are analyzed under four different hypotheses:

H_1: $\beta_1 = \beta_2$, $\beta_3 = \beta_4$, $\theta_1^2 = \theta_2^2$, $\theta_3^2 = \theta_4^2$, $\rho = 1$

H_2: $\beta_1 = \beta_2$, $\beta_3 = \beta_4$, $\theta_1^2 = \theta_2^2$, $\theta_3^2 = \theta_4^2$

H_3: $\rho = 1$

H_4: Σ is of the form (15) with β_1, β_2, β_3, β_4, θ_1, θ_2, θ_3, θ_4, and ρ unconstrained.

The results are shown in Table 2b. Each hypothesis is tested against the general alternative that Σ is unconstrained. To consider various hypotheses that can be tested, the four χ^2 values of Table 2b are recorded in a 2 × 2 table as in Table 2c. Test of H_1 against H_2 gives $\chi^2 = 35.40$ with 1 degree of freedom. An alternative test is H_3 against H_4, which gives $\chi^2 = 35.51$ with 1 degree of freedom. Thus, regardless of whether we treat the two pairs of tests as parallel or congeneric, the hypothesis $\rho = 1$ is rejected. There is strong evidence that the unspeeded and speeded tests do not measure the same trait. I shall return to this question in the next section. The hypothesis of parallelness of the two pairs of tests can also be tested by means of Table 2c. This gives $\chi^2 = 1.12$ or $\chi^2 = 1.23$ with 4 degrees of freedom, depending on whether we assume $\rho = 1$ or $\rho \neq 1$. Thus we cannot reject the hypothesis that the two pairs of tests are parallel. It appears that H_2 is the most reasonable of the four hypotheses. The maximum likelihood estimate of ρ under H_2 is $\hat{\rho} = 0.899$ with a standard error of 0.019. An approximate 95% confidence interval for ρ is $0.86 < \rho < 0.94$.

TABLE 2

Lord's Vocabulary Test Data
Covariance Matrix (a)

N = 649

	x_1	x_2	y_1	y_2
x_1	86.3979			
x_2	57.7751	86.2632		
y_1	56.8651	59.3177	97.2850	
y_2	58.8986	59.6683	73.8201	97.8192

Summary of Analyses (b)

Hypothesis	No. par.	χ^2	d.f.	P
H_1	4	37.33	6	0.00
H_2	5	1.93	5	0.86
H_3	8	36.21	2	0.00
H_4	9	0.70	1	0.40

Tests of Hypotheses (c)

	Parallel	Congeneric	
$\rho = 1$	$\chi_6^2 = 37.33$	$\chi_2^2 = 36.21$	$\chi_4^2 = 1.12$
$\rho \neq 1$	$\chi_5^2 = 1.93$	$\chi_1^2 = 0.70$	$\chi_4^2 = 1.23$
	$\chi_1^2 = 35.40$	$\chi_1^2 = 35.51$	

Analysis of Speeded and Unspeeded Tests

Table 3a shows correlations and standard deviations for 18 tests. There are three kinds of tests, Vocabulary, Intersections, and Arithmetic Reasoning tests and of each kind there are two level tests (L) which are unspeeded or very little speeded, one medium-speeded (M) and three highly speeded (S). These tests together with several others were analyzed by Lord [1956] with the objective to isolate and identify speed factors in tests and their relationships to academic grades. Lord (p. 31) lists among others the following questions, that his study was designed to answer: "Is speed on cognitive tests a unitary trait? Or are there different kinds of speed for different kinds of tasks? If so, how highly correlated are these kinds of speed? How highly correlated are speed and level on the same task?" Lord used maximum likelihood factor analysis and oblique rotations, and verbal-speed and spatial-speed factors were clearly identified but no arithmetic-reasoning speed factor was found. A smaller battery of 15 tests was used by Jöreskog [1971] to answer the same kinds of questions by means of a series of confirmatory analyses. One of these models will be used here.

I postulate that there are three power factors and three speed factors, one of each kind for each kind of task. The two unspeeded tests load only on the power factor whereas the four speeded tests load on both the power and the speed factor. The speed factors are postulated to be uncorrelated with the power factors. The resulting solution, given in Table 3b, has $\chi^2 = 167.51$ with 117 degrees of freedom. This has a P-level of 0.002. Table 3b also gives approximate 95% confidence intervals for all the estimated parameters. Note that the loadings for the medium speeded tests on the speed factor are on the boundary of being significant. In order to detect and isolate the speed factor in these medium speeded tests one should either have a larger sample or these tests should be more speeded. The three power factors correlate moderately and so do the three speed factors. Thus all six factors are clearly manifested. The advantage of postulating that speed and power factors are uncorrelated is that one can easily get a variance decomposition of each total test variance into components due to power, speed and error. These variance components are obtained simply by squaring the corresponding loadings in Table 3b.

Can the fit of the model in Table 3b be improved by letting speed and power factors correlate? The answer is no for the following reason. If the correlation between factors V and VS, say, is relaxed, one can add a multiple of column 4 of the factor matrix to column 1. This change can be compensated by changing the correlation between V and VS in such a way that Σ remains the same. Thus, the correlation between V and VS and the loadings of the four speeded tests on the power factor are not identified while all the other parameters are identified. In fact, the correlation between V and VS can be assigned any value (between zero and one) and the fit of the model will be identically the same. Although one cannot estimate the correlations between speed and power on the same task, one can estimate the correlations between speed and power factors on different tasks. However, these correlations are generally very small and they do not improve fit significantly.

4. Factor Analysis Models

The models of the previous section are all special cases of a general model for confirmatory factor analysis developed by Jöreskog [1969]. However, factor analysis is most widely used in exploratory studies, in which the nature of the common factors is to be revealed by the analysis rather than postulated in advance. The basic idea is that for a given set of response variates $x_1, \cdots, x_p$ one wants to find a set of underlying or latent factors $f_1, \cdots, f_k$, fewer in number than the observed variates, that will account for the intercorrelations of the response variates, in the sense that when the factors are partialled

TABLE 3a

Lord's Speed Factor Data

Total Sample N = 649

		Means	Standard Deviations	Correlations																		
Vocabulary	L	49.49	9.30	1.00																		
"	L	49.45	9.29	.67	1.00																	
"	M	49.63	9.73	.71	.69	1.00																
"	S	49.76	9.86	.62	.65	.66	1.00															
"	S	49.90	9.61	.69	.70	.74	.77	1.00														
"	S	49.83	9.89	.64	.65	.70	.76	.85	1.00													
Intersections	L	49.94	9.55	.15	.18	.11	.16	.12	.09	1.00												
"	L	50.15	9.67	.06	.14	.08	.14	.06	.07	.72	1.00											
"	M	49.97	9.81	.10	.18	.12	.18	.10	.12	.71	.77	1.00											
"	S	50.00	9.80	.12	.17	.09	.19	.13	.14	.70	.71	.75	1.00										
"	S	50.10	9.83	.09	.13	.08	.17	.11	.10	.70	.74	.76	.80	1.00									
"	S	50.07	9.88	.08	.12	.08	.16	.10	.11	.70	.75	.77	.79	.83	1.00								
Arith. Reasoning	L	49.52	9.05	.30	.32	.27	.28	.29	.27	.34	.30	.34	.32	.31	.33	1.00						
"	L	50.16	9.67	.25	.31	.26	.29	.27	.29	.36	.30	.36	.30	.29	.31	.54	1.00					
"	M	49.92	9.73	.29	.34	.30	.32	.31	.35	.30	.30	.34	.31	.29	.32	.61	.56	1.00				
"	S	49.79	9.91	.28	.26	.30	.36	.35	.38	.29	.28	.33	.31	.29	.36	.55	.54	.63	1.00			
"	S	49.80	9.85	.27	.29	.30	.33	.35	.35	.34	.28	.34	.35	.34	.37	.53	.55	.58	.63	1.00		
"	S	50.00	9.83	.28	.30	.31	.38	.36	.39	.26	.23	.25	.27	.25	.31	.51	.53	.57	.64	.61	1.00

TABLE 3b

Lord's Speed Factor Data

Total Sample N = 649

		V	S	R	VS	SS	RS	Unique Variance
Vocabulary	L	82 ± 07	00*	00*	00*	00*	00*	33 ± 04
"	L	83 ± 07	00*	00*	00*	00*	00*	31 ± 04
"	M	84 ± 07	00*	00*	09 ± 07	00*	00*	29 ± 04
"	S	76 ± 08	00*	00*	33 ± 08	00*	00*	31 ± 04
"	S	85 ± 07	00*	00*	40 ± 07	00*	00*	13 ± 03
"	S	79 ± 07	00*	00*	48 ± 07	00*	00*	15 ± 04
Intersections	L	00*	83 ± 07	00*	00*	00*	00*	32 ± 04
"	L	00*	87 ± 07	00*	00*	00*	00*	24 ± 04
"	M	00*	88 ± 07	00*	00*	06 ± 07	00*	23 ± 04
"	S	00*	84 ± 07	00*	00*	25 ± 10	00*	24 ± 03
"	S	00*	85 ± 07	00*	00*	32 ± 07	00*	18 ± 03
"	S	00*	86 ± 06	00*	00*	34 ± 10	00*	16 ± 03
Arith. Reasoning	L	00*	00*	75 ± 06	00*	00*	00*	44 ± 06
"	L	00*	00*	74 ± 09	00*	00*	00*	46 ± 06
"	M	00*	00*	79 ± 07	00*	00*	11 ± 07	38 ± 05
"	S	00*	00*	74 ± 08	00*	00*	36 ± 09	32 ± 05
"	S	00*	00*	72 ± 07	00*	00*	29 ± 07	40 ± 05
"	S	00*	00*	70 ± 08	00*	00*	35 ± 10	39 ± 06

	V	S	R	VS	SS	RS
V	100*					
S	16 ± 08	100*				
R	49 ± 07	48 ± 07	100*			
VS	00*	00*	00*	100*		
SS	00*	00*	00*	27 ± 16	100*	
RS	00*	00*	00*	65 ± 18	41 ± 22	100*

$\chi^2 = 167.51$ with 117 degrees of freedom

P =   0.002

Asterisks denote parameter values specified by hypothesis.

All entries have been multiplied by 100.

out from the observed variates there no longer remains any correlation between these. This leads to the model

(16) $x = \mu + \Lambda f + z$

where $\mathcal{E}(x) = \mu$, $\mathcal{E}(f) = 0$ and $\mathcal{E}(z) = 0$, z being uncorrelated with f. Let $\Phi = \mathcal{E}(ff')$ which may be taken as a correlation matrix and $\Psi^2 = \mathcal{E}(zz')$ which is diagonal. Then the covariance matrix Σ of x becomes

(17) $\Sigma = \Lambda \Phi \Lambda' + \Psi^2$.

If $(p - k)^2 < p + k$, this relationship can be tested statistically, unlike (16) which involves hypothetical variates and cannot be verified directly.

When $k > 1$ there is an indeterminacy in (16) arising from the fact that a nonsingular linear transformation of f changes Λ and in general also Φ but leaves Σ unchanged. The usual way to deal with this indeterminacy in exploratory factor analysis [see, e.g., Lawley & Maxwell, 1971 or Jöreskog, 1967] is to choose $\Phi = I$ and $\Lambda'\Psi^{-2}\Lambda$ or $\Lambda'\Lambda$ to be diagonal and to estimate the parameters in Λ and Ψ subject to these conditions. This leads to an arbitrary set of factors which may then be subjected to a rotation or a linear

transformation to another set of factors which can be given a more meaningful inter-
pretation. This rotation is usually guided by Thurstone's [1947] principle of simple
structure which states that in each column of Λ there should only be a relatively small
number of large loadings and a relatively large number of small loadings. Efficient
computational procedures for estimating the unrotated factor loadings by ULS, GLS or
ML have been described by Jöreskog [1977a] and a computer program by Jöreskog and
Sörbom [1979] is available for this. Analytical procedures for rotation to simple structure
are described by Harman [1967].

The factor analyst usually only interprets the large loadings in a factor matrix Λ after
rotation to simple structure. However, the most difficult question which have plagued
factor analysts for several decades is "How large should a factor loading be to be
considered significant?" or "How small should a factor loading be to be ignored?". The
first break-through came with the papers by Archer and Jennrich [1973] and Jennrich
[1973] who developed procedures for estimating standard errors for rotated factor load-
ings estimated by the maximum likelihood method. Unfortunately, however, the amount
of computation involved in these procedures is so enormous that their feasibility in
practice becomes very limited. I shall therefore propose an alternative procedure here.
This procedure not only gives standard errors for factor loadings but, when applied
routinely, also yields what I shall call "the best-fitting simple structure" for a given
number of factors.

Finding the Best Fitting Simple Structure

Once a sufficient number of restrictions has been imposed on the model to make it
identified, standard errors for factor loadings can be estimated by the method described in
Section 2. The simplest way to achieve identification, assuming that Φ is a correlation
matrix with one's in the diagonal, is to set at least $k - 1$ zeroes in each column of Λ.
Sometimes one has enough knowledge about the factorial nature of the tests to be able to
specify a priori that certain variables should not load on certain factors. If this is not
possible, as in a completely exploratory analysis, one can proceed as follows:
  (i)   rotate the factors orthogonally by the varimax method [Kaiser, 1958];
  (ii) raise the varimax factor loadings to some power 3 or 4 while retaining their signs.
        Use these numbers as a target and perform a promax [Hendrickson and White,
        1964] procrustes rotation. This produces an oblique solution.
  (iii) in the promax factor matrix find the largest factor loading in each column and,
        assuming that these are in different rows, rotate the original unrotated factor
        matrix or the varimax factor matrix so that the other loadings in these rows
        become zero. This results in an oblique solution in which each factor vector
        passes through one test point.
The solution obtained in (iii) has $k - 1$ zeroes in each column of Λ and has exactly
the same fit to the data as any other rotated solution. A confirmatory analysis with the
same fixed zeroes will yield the same non-zero factor loadings but standard errors for these
can now be obtained. An inspection of the loadings in relation to their standard errors will
usually reveal that a large number of the loadings are insignificant. By eliminating these,
i.e. by setting them equal to zero, one can reduce a lot of noise in the model and estimate
the really significant loadings more precisely. To obtain the best-fitting simple structure,
proceed with the next step:
  (iv) set all insignificant loadings to zero. The resulting solution will not in general
        have a significantly worse fit than the original solution but will display a neat
        simple structure with many zero loadings.
Since it is a possibility that the loadings that were set to zero in step (iii) are not zero,
it is recommended as a safeguarding final step that the zero loadings are checked again as

follows. Find the largest absolute derivative of F with respect to the fixed zero loadings and relax this particular zero loading while keeping all other zeroes fixed. This loading will make the function decrease maximally and hence make the largest improvement in fit. If the improvement in fit is significant this step should be repeated again. Otherwise the final solution is "best-fitting" in the following sense:

(a) all non-zero loadings are significant;

(b) all zero loadings are such that if they were relaxed they would not be significant.

The whole procedure for finding the best fitting simple structure will be illustrated using the well-known Harman's [1967] twenty-four psychological variables with four factors. The varimax solution is given in Table 4a. This has an overall χ^2 of 246.36 with 186 degrees of freedom representing a reasonably good fit of the unrestricted four-factor model. Since this already exhibits a rather good simple structure I shall omit the step (ii) of the procedure and continue directly with step (iii). The largest loading in each column is underlined in Table 4a. Using the tests 9, 10, 1 and 17 as reference variables for Factors 1, 2, 3 and 4, respectively, one obtains the solution given in Table 4b in which the standard errors of the non-zero loadings are given in parenthesis. Setting to zero those loadings which are not significant at the 5% level, i.e. those that are less in magnitude than two times its standard error, gives the solution in Table 4c. This has a χ^2 of 301.42 with 231 degrees of freedom. The difference between this and the original χ^2 is 55.06 with 45 degrees of freedom which is not significant. Hence, although the solution in Table 4c represents a highly restricted solution with many zero loadings, the fit of this solution is not any worse than that of the original unrestricted solution. The largest derivative occurs for the zero loading in row 12 and column 1 of Table 4c. When this loading is relaxed, χ^2 drops to 295.52, a drop of 5.89 with one degree of freedom. Since this is not significant at the 1% level the zero loading is retained and the solution in Table 4c is accepted as the best fitting simple structure.

The interpretation of the factors in Table 4c is straightforward. The first factor is a logical verbal reasoning factor, the second is a speed factor, the third is a spatial visualization factor and the fourth is a memory factor.

5. Variance and Covariance Components

Estimation of Variance Components

Several authors [Bock, 1960; Bock & Bargmann, 1966 and Wiley, Schmidt & Bramble, 1971] have considered covariance structure analysis as an approach to study differences in test performance when the tests have been constructed by assigning items or subtests according to objective features of content or format to subclasses of a factorial or hierarchical classification.

Bock [1960] suggested that the scores of N subjects on a set of tests classified in 2^n factorial design may be viewed as data from an $N \times 2^n$ experimental design, where the subjects represent a random mode of classification and the tests represent n fixed mode of classification. Bock pointed out that conventional mixed-model analysis of variance gives useful information about the psychometric properties of the tests. In particular, the presence of non-zero variance components for the random mode of classification and for the interaction of the random and fixed modes of classification provides information about the number of dimensions in which the tests are able to discriminate among subjects. The relative size of these components measure the power of the tests to discriminate among subjects along the respective dimensions.

The multitrait-multimethod matrix of Campbell and Fiske [1959] is an example of a factorial design of tests. A more complex design is Guilford's [1956] structure of intellect.

TABLE 4a

Twentyfour Psychological Variables

Varimax-Rotated ML- Solution

Reference factor loadings underlined

	1	2	3	4
1. Visual Perception	0.160	0.187	<u>0.689</u>	0.160
2. Cubes	0.117	0.083	0.436	0.096
3. Paper Form Board	0.137	-0.019	0.570	0.110
4. Flags	0.233	0.099	0.527	0.080
5. General Information	0.739	0.213	0.185	0.150
6. Paragraph Comprehension	0.767	0.066	0.205	0.233
7. Sentence Completion	0.806	0.153	0.197	0.075
8. Word Classification	0.569	0.242	0.338	0.132
9. Word Meaning	<u>0.806</u>	0.040	0.201	0.227
10. Addition	0.168	<u>0.831</u>	-0.118	0.167
11. Code	0.180	0.512	0.120	0.374
12. Counting Dots	0.019	0.716	0.210	0.088
13. Straight-Curved Capitals	0.188	0.525	0.438	0.082
14. Word Recognition	0.197	0.081	0.050	0.553
15. Number Recognition	0.122	0.074	0.116	0.520
16. Figure Recognition	0.069	0.062	0.408	0.525
17. Object-Number	0.142	0.219	0.062	<u>0.574</u>
18. Number-Figure	0.026	0.336	0.293	0.456
19. Figure-Word	0.148	0.161	0.239	0.365
20. Deduction	0.378	0.118	0.402	0.301
21. Numerical Puzzles	0.175	0.438	0.381	0.223
22. Problem Reasoning	0.366	0.122	0.399	0.301
23. Series Completion	0.369	0.244	0.500	0.239
24. Arithmetic Problems	0.370	0.496	0.157	0.304

$\chi^2 = 246.36$ with 186 d.f.

This design is based on a cross-classification of test items, not all of which may exist or be employed in any one study. Thus the classification scheme may be incomplete.

Consider an experimental design that has one random way of classification $\nu = 1, 2, \ldots, N$, one fixed way of classification $i = 1, 2, 3$ and another fixed way of classification $j =$

TABLE 4b

Twentyfour Psychological Variables

Reference-Variables ML-Solution

Obtained by oblique transformation of varimax solution

Standard errors in parenthesis

Factor Loadings

	1	2	3	4
1. Visual Perception	-0.000	-0.000	0.749(0.078)	0.000
2. Cubes	0.027(0.110)	-0.034(0.111)	0.462(0.111)	-0.004(0.138)
3. Paper Form Board	0.049(0.108)	-0.177(0.109)	0.580(0.108)	0.008(0.136)
4. Flags	0.148(0.107)	-0.017(0.108)	0.541(0.107)	-0.079(0.135)
5. General Information	0.762(0.090)	0.214(0.088)	0.052(0.090)	-0.113(0.108)
6. Paragraph Comprehension	0.809(0.087)	0.020(0.083)	0.019(0.087)	0.013(0.104)
7. Sentence Completion	0.866(0.094)	0.186(0.091)	0.045(0.093)	-0.219(0.110)
8. Word Classification	0.533(0.094)	0.198(0.094)	0.274(0.094)	-0.116(0.116)
9. Word Meaning	0.862(0.068)	-0.000	0.000	0.000
10. Addition	-0.000	0.872(0.076)	-0.000	0.000
11. Code	0.007(0.097)	0.391(0.099)	0.155(0.101)	0.292(0.118)
12. Counting Dots	-0.190(0.097)	0.690(0.104)	0.387(0.098)	-0.085(0.122)
13. Straight-Curved Capitals	0.014(0.103)	0.456(0.106)	0.555(0.103)	-0.141(0.130)
14. Word Recognition	0.102(0.125)	-0.131(0.129)	-0.057(0.132)	0.622(0.148)
15. Number Recognition	0.010(0.125)	-0.148(0.129)	0.039(0.131)	0.593(0.147)
16. Figure Recognition	-0.100(0.123)	-0.235(0.126)	0.374(0.124)	0.570(0.143)
17. Object-Number	0.000	-0.000	0.000	0.634(0.086)
18. Number-Figure	-0.179(0.114)	0.110(0.115)	0.332(0.115)	0.454(0.134)
19. Figure-Word	0.029(0.111)	-0.014(0.113)	0.213(0.114)	0.347(0.135)
20. Deduction	0.294(0.098)	-0.043(0.099)	0.337(0.099)	0.176(0.121)
21. Numerical Puzzles	0.002(0.100)	0.315(0.101)	0.452(0.100)	0.072(0.125)
22. Problem Reasoning	0.279(0.099)	-0.039(0.100)	0.338(0.100)	0.180(0.122)
23. Series Completion	0.250(0.095)	0.096(0.096)	0.488(0.094)	0.056(0.120)
24. Arithmetic Problems	0.240(0.089)	0.417(0.091)	0.156(0.092)	0.141(0.110)

Factor Intercorrelations

	1	2	3	4
1	1.000			
2	0.243(0.106)	1.000		
3	0.482(0.108)	0.196(0.124)	1.000	
4	0.487(0.127)	0.533(0.126)	0.418(0.153)	1.000

1, 2, 3 for $i = 1$, 2 and $j = 1$, 2 for $i = 3$. One model that may be considered is

(18) $$x_{vij} = \mu_{ij} + a_v + b_{vi} + c_{vj} + e_{vij},$$

where μ_{ij} is the mean of x_{vij} and where a_v, b_{vi}, c_{vj} and e_{vij} are uncorrelated random variables with zero means and variances σ_a^2, $\sigma_{b_i}^2$, $\sigma_{c_j}^2$ and $\sigma_{e_{ij}}^2$, respectively. Writing $\mathbf{x}_v' = (x_{v11}, x_{v12}, x_{v13}, x_{v21}, x_{v22}, x_{v23}, x_{v31}, x_{v32})$, $\mathbf{u}_v' = (a_v, b_{v1}, b_{v2}, b_{v3}, c_{v1}, c_{v2}, c_{v3})$ and

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix},$$

we may write (18) as

$$\mathbf{x}_v = \mathbf{\mu} + \mathbf{A}\mathbf{u}_v + \mathbf{e}_v$$

where $\mathbf{\mu}$ is the mean vector and $\mathbf{e}$ is a random error vector both of the same form as $\mathbf{x}_v$. The

TABLE 4 c.

Twentyfour Psychological Variables.

Best-Fitting Simple Structure ML-Solution

Factor Loadings

	1	2	3	4
1. Visual Perception	0.0	0.0	0.736	0.0
2. Cubes	0.0	0.0	0.468	0.0
3. Paper Form Board	0.0	0.0	0.553	0.0
4. Flags	0.0	0.0	0.594	0.0
5. General Information	0.758	0.148	0.0	0.0
6. Paragraph Comprehension	0.824	0.0	0.0	0.0
7. Sentence Completion	0.806	0.070	0.0	0.0
8. Word Classification	0.512	0.158	0.219	0.0
9. Word Meaning	0.868	0.0	0.0	0.0
10. Addition	0.0	0.860	0.0	0.0
11. Code	0.0	0.412	0.0	0.394
12. Counting Dots	0.0	0.642	0.209	0.0
13. Straight-Curved Capitals	0.0	0.435	0.452	0.0
14. Word Recognition	0.0	0.0	0.0	0.547
15. Number Recognition	0.0	0.0	0.0	0.529
16. Figure Recognition	0.0	0.0	0.338	0.375
17. Object-Number	0.0	0.0	0.0	0.649
18. Number-Figure	0.0	0.0	0.172	0.510
19. Figure-Word	0.0	0.0	0.0	0.493
20. Deduction	0.324	0.0	0.414	0.0
21. Numerical Puzzle	0.0	0.383	0.463	0.0
22. Problem Reasoning	0.320	0.0	0.403	0.0
23. Series Completion	0.258	0.146	0.503	0.0
24. Arithmetic Problems	0.384	0.498	0.0	0.0

Factor Intercorrelations

	1	2	3	4
1	1.000			
2	0.234	1.000		
3	0.505	0.200	1.000	
4	0.476	0.440	0.520	1.000

$$\chi^2 = 301.42 \quad \text{with} \quad 231 \text{ d.f.}$$

covariance matrix of $\mathbf{x}_\nu$ is

$$(19) \qquad\qquad \mathbf{\Sigma} = \mathbf{A\Phi A'} + \mathbf{\Psi'^2}$$

where $\mathbf{\Phi}$ is a diagonal matrix whose diagonal elements are σ^2_a, $\sigma^2_{b_1}$, $\sigma^2_{b_2}$, $\sigma^2_{b_3}$, $\sigma^2_{c_1}$, $\sigma^2_{c_2}$ and $\sigma^2_{a_3}$ and $\mathbf{\Psi'^2}$ is a diagonal matrix whose elements are the $\sigma^2_{e_{ij}}$. The matrix $\mathbf{A}$ has rank 5, and only 5 linearly independent combinations of the components of $\mathbf{u}_\nu$ are estimable [see e.g. Graybill, 1961, pp 228–229]. In conventional mixed-model analysis of variance one usually makes the assumptions that $\sigma^2_{b_i} = \sigma^2_b$ for $i = 1, 2, 3$, $\sigma^2_{c_j} = \sigma^2_c$ for $j = 1, 2, 3$ and $\sigma^2_{e_{ij}} = \sigma^2_e$ for all i and j, but all these assumptions are not necessary.

In general, if $\mathbf{A}$ is of order $p \times r$ and of rank k, one may choose k independent linear functions, each one linearly dependent on the rows of $\mathbf{A}$ and estimate the covariance matrix of these functions. It is customary to choose linear combinations that are mutually uncorrelated but this is not necessary. Let $\mathbf{L}$ be the matrix of coefficients of the chosen linear functions and let $\mathbf{K}$ be any matrix such that $\mathbf{A} = \mathbf{KL}$. For example, $\mathbf{K}$ may be obtained from

$$(20) \qquad\qquad \mathbf{K} = \mathbf{AL'(LL')^{-1}}.$$

The model may then be reparameterized to full rank by defining $\mathbf{u^*} = \mathbf{Lu}$. We then have $\mathbf{x} = \mathbf{Au} + \mathbf{e} = \mathbf{KLu} + \mathbf{e} = \mathbf{Ku^*} + \mathbf{e}$. The covariance matrix of $\mathbf{x}$ is represented as

$$(21) \qquad\qquad \mathbf{\Sigma} = \mathbf{K\Phi^* K'} + \mathbf{\Psi'^2}$$

where $\mathbf{\Phi^*}$ is the covariance matrix of $\mathbf{u^*}$ and $\mathbf{\Psi'^2}$ is as before. One can now estimate $\mathbf{\Psi'^2}$ and $\mathbf{\Phi^*}$. The latter may be taken to be diagonal if desired.

It should be noted that the covariance structure in (21) is linear in $\mathbf{\Phi^*}$ and $\mathbf{\Psi'^2}$. When the elements of $\mathbf{\Phi^*}$ and $\mathbf{\Psi'^2}$ are considered as the parameters generating $\mathbf{\Sigma}$ then $\partial\mathbf{\Sigma}/\partial\theta_i$ in (8) are constant and independent of $\mathbf{\theta}$ and as a consequence the information matrix in (8) is constant in ULS and GLS. Browne [1974] used this fact and showed how GLS estimators can be obtained in closed form. With the algorithm outlined in Section 2, the ULS and GLS estimators are obtained after one iteration since the function being minimized is exactly quadratic and the information matrix in (8) is the exact Hessian, which is constant.

A General Class of Components of Covariance Models

The models of the previous subsection assume that all tests or subtests are measured on the same scale. Wiley, Schmidt and Bramble [1973] suggested the study of a general class of components of covariance models which would allow different tests to be on different scales. The covariance matrix $\mathbf{\Sigma}$ will then be of the form

$$(22\text{a-b}) \qquad \mathbf{\Sigma} = \mathbf{\Delta A\Phi A'\Delta} + \mathbf{\Theta^2} \text{ or } \mathbf{\Sigma} = \mathbf{\Delta(A\Phi A'} + \mathbf{\Psi'^2)\Delta}.$$

The matrix $\mathbf{A}(p \times k)$ is assumed to be known and gives the coefficient of the linear functions connecting the manifest and latent variables, $\mathbf{\Delta}$ is a $p \times p$ diagonal matrix of unknown scale factors, $\mathbf{\Phi}$ is the $k \times k$ symmetric and positive definite covariance matrix of the latent variables and $\mathbf{\Psi'^2}$, and $\mathbf{\Theta^2}$ are $p \times p$ diagonal matrices of error variances.

Within this class of models eight different special cases are of interest. These are generated by the combination of the following set of conditions;

$$\text{on } \mathbf{\Delta} : \left\{ \begin{array}{l} \mathbf{\Delta = I} \\ \mathbf{\Delta \neq I} \end{array} \right\}$$

$$\text{on } \mathbf{\Phi} : \left\{ \begin{array}{l} \mathbf{\Phi} \text{ is diagonal} \\ \mathbf{\Phi} \text{ is not diagonal} \end{array} \right\}$$

$$\text{on } \Psi'^2 \text{ or } \Theta^2 : \left\{ \begin{array}{c} \Psi'^2 \text{ or } \Theta^2 = \sigma^2 \, \mathbf{I} \\ \Psi'^2 \text{ or } \Theta^2 \text{ general diagonal} \end{array} \right\}.$$

The classical formulation of the mixed model and its generalizations assume that $\Delta = \mathbf{I}$. This is appropriate if the observed variables are in the same metric as for example when the observed variables represent physical measurements, time to criterion measures, reaction times or items similarly scaled such as semantic differential responses. However, if the observed variables are measured in different metrics then the classical model would not fit. In such cases the inclusion of Δ in the model as a general diagonal matrix of scaling factors would provide a useful alternative specification. It should be pointed out that the elements of Δ do not have to be related to the variances of the variables.

The classical components of the variance model assume that Φ is diagonal. However, there are usually no substantive reasons for assuming this.

The two conditions on Ψ'^2 or Θ^2 correspond to homogeneous and heterogeneous error variances. If the variables are in the same metric and if the measurement situation is sufficiently similar from variable to variable then it would seem reasonable to hypothesize that the variances of the errors of measurement ought to be homogeneous, i.e., in (22a) we take $\Delta = \mathbf{I}$ and $\Theta^2 = \sigma^2 \mathbf{I}$.

If, on the other hand, the scale of measurement is the same but the measurement situation from variable to variable is different enough to generate different kinds of error structures, then the variances of the errors of measurement might differ systematically from variable to variable. For this situation it would seem best to take $\Delta = \mathbf{I}$ but leave Θ^2 free in (22a). If the manifest variables were in different metrics then clearly the error variances in the observed metric will most likely be heterogeneous. One useful hypothesis to test in this context would be that the standard deviations of the errors of measurement are proportional to the rescaling factors. This would correspond to taking $\Psi'^2 = \sigma^2 \mathbf{I}$ in (22b). When both Δ and Ψ'^2 are free, (22a) and (22b) are equivalent.

Analysis of a 2^3 Factorial Design

This example is taken from Wiley, Schmidt and Bramble [1973]. The original data are from a study by Miller and Lutz [1966], and consist of the scores of 51 education students on a test designed to assess teachers' judgments concerning the effects of situation and instruction factors on the facilitation of pupil learning. The items used here were designed according to three factors which were hypothesized to influence classroom learning situations and teaching practices. The three factors and their levels are given as [see Miller & Lutz, 1966]:

Grade Level (G). The levels of this factor, the first grade (G_1) and the sixth grade (G_2), were chosen to represent extremes of the elementary grades. In this way it was possible to maximize the opportunity for observing any differences in teachers' judgments that might occur as a result of variations due to grade level.

Teacher Approach (T). The teacher-centered approach (T_1) and the pupil-centered approach (T_2) were distinguished as levels of this factor on the basis of the locus of described activity control, and the direction of described pupils' attention. In the case of the teacher-centered approach, the locus and direction were oriented to the teacher; in the case of the pupil-centered approach, the locus and direction were oriented to the pupil.

Teaching Method (M). Level one of this factor was drill (M_1) which was used strictly to refer to rote learning activities; discovery (M_2) was used to refer to an approach in which the teacher attempts to develop pupil understanding through procedures aimed at simulating insight without recourse to rote memorization or rigid learning routines.

The eight subtest scores are each based on eight items. The eight subtests conform to

a 2^3 factorial arrangement and is given by

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & -1 \end{bmatrix}.$$

The sample covariance matrix $\mathbf{S}$ for these data is given in Table 5a.

For this data the use of $\mathbf{\Delta} = \mathbf{I}$ was considered appropriate because the same scale (1 − 7) was employed for each item and the numbers of items in the subtests were also equal.

The data was analyzed under each of the four remaining cases in the general class of models considered in the previous section. The χ^2 values and their degrees of freedom are shown in Table 5b. In the margin of this table, χ^2-values are given for testing the hypotheses of uncorrelated components and of homogeneous error variances. It appears that these hypotheses are both rejected by the tests. The only model that seems reasonable is the one which assumes both correlated components and heterogenous error variances. The maximum likelihood estimates of the variances and covariances of the components and the error variances, together with their standard errors are given in Table 5c.

The relative magnitudes of the estimated variance components for the latent variables indicates the major sources of variation in the performance of the subjects. The estimate of the first variance component ($\hat{\phi}_{11} = 9.16$) is the largest as would be expected since this component reflects the variation due to individual differences between the subjects. The estimated values of the other components indicate that another major source of variation in the responses of the subjects is due to the different type of teaching method specified in the content of the item (i.e., $\hat{\phi}_{44} = 5.21$). Apparently the education students varied considerably as to how the contrast between drill and discovery methods of instruction influenced their responses. Variation contributed by grade level was intermediate in magnitude ($\hat{\phi}_{22} = .70$). The estimate of the variance component for the teacher approach factor and its large standard error (i.e., $\hat{\phi}_{33} = .43$, S.E. ($\hat{\phi}_{33}$) = .91) indicate that this was not an important source of variation in the performance of the subjects. One of the estimated covariances was relatively large—that between the teaching method factor and the teacher approach factor—indicating these latent variables to be highly correlated. This would indicate that the responses of the education students to the different type of teaching method specified in the items was related to their responses to the teacher approach factor found in the items.

6. Measurement Errors in Regression Models

The models of Sections 3–5 are all such that all observed variables are of the same kind; they are all generated by some latent variables of one kind or another. In this section I consider models in which there are two kinds of observed variables, dependent or caused variables denoted by y and independent or causal variables denoted by x. The purpose of the models is to account for, or estimate, or predict the variables y. The variables x are given variables, usually freely correlated, which are believed to influence the y-variables. These x-variables may be random or fixed variables.

The ordinary linear regression of one y-variable on a number of x-variables is

(23) $$y = \gamma' x + z$$

TABLE 5

Miller-Lutz Data

(taken from Wiley, Schmidt and Bramble, 1973)

a. Sample Covariance Matrix

$$
\underset{\sim}{S} = \begin{bmatrix}
18.74 & & & & & & & \\
9.28 & 18.80 & & & & & & \\
15.51 & 7.32 & 21.93 & & \text{(symmetric)} & & & \\
3.98 & 15.27 & 4.10 & 26.62 & & & & \\
15.94 & 4.58 & 13.79 & -2.63 & 19.82 & & & \\
7.15 & 13.63 & 3.86 & 15.33 & 3.65 & 16.81 & & \\
11.69 & 6.05 & 10.18 & 1.13 & 13.55 & 5.72 & 16.58 & \\
2.49 & 12.35 & .03 & 16.93 & -.86 & 14.33 & 2.99 & 18.26
\end{bmatrix}
$$

b. χ^2 Values for Testing the Fit of Four Models

	$\underset{\sim}{\Theta}^2 = \sigma^2 \underset{\sim}{I}$	$\underset{\sim}{\Theta}^2 \neq \sigma^2 \underset{\sim}{I}$	
$\underset{\sim}{\Phi}$ diagonal	$\chi^2_{31} = 68.25$	$\chi^2_{24} = 46.16$	$\chi^2_7 = 22.09$
$\underset{\sim}{\Phi}$ not diagonal	$\chi^2_{25} = 51.00$	$\chi^2_{18} = 25.98$	$\chi^2_7 = 25.02$
	$\chi^2_6 = 17.25$	$\chi^2_6 = 20.18$	

c. Maximum Likelihood Estimates of Φ and Θ^2

Standard Errors in Parenthesis

$$
\hat{\underset{\sim}{\Phi}} = \begin{bmatrix}
9.16(1.95) & & & \\
0.75(0.48) & 0.70(0.34) & & \\
0.63(0.43) & -0.05(0.33) & 0.43(0.91) & \\
-0.62(1.10) & -0.51(0.81) & 1.13(0.51) & 5.21(1.58)
\end{bmatrix}
$$

$\hat{\underset{\sim}{\Theta}}^2 = \text{diag}\ [1.52(0.83),\ 4.95(1.41),\ 8.25(1.88),\ 5.58(1.60),\ 1.95(0.96),$
    $5.76(1.21),\ 2.52(0.92)]$

where the residual or random disturbance term z is assumed to be uncorrelated with $\mathbf{x}$. The covariance structure for this model is

$$(24) \qquad \Sigma = \begin{pmatrix} \gamma' \Sigma_{xx} \gamma + \sigma_z^2 & \\ \Sigma_{xx} \gamma & \Sigma_{xx} \end{pmatrix}.$$

If x is fixed, Σ_{xx} is S_{xx}, the covariance matrix computed from the fixed values of $\mathbf{x}$. If $\mathbf{x}$ is random the estimate of Σ_{xx} is S_{xx}. In both cases, γ is estimated from the covariances in the first column of S and σ_z^2 is estimated as $s_{yy} - \hat{\gamma}' S_{xx} \hat{\gamma}$. There is a one to one correspondence between the parameters Σ_{xx}, γ and σ_z^2 of the regression model and the variances and covariances Σ_{xx}, δ_{xy} and σ_{yy} in Σ. The model is just identified and ULS, GLS and ML yield the same estimators all with a perfect fit.

When there is measurement error in one or more of the x-variables or in the y-variable and one is interested in the true regression

$$(25) \qquad \eta = \gamma'\xi + \zeta$$

rather than (23), then the above estimator of γ is not a consistent estimator of the γ in (25). Here ζ is the residual in the true regression (25), assumed to be uncorrelated with ξ, and y and x are observed variables having η and ξ as common factors or true scores, i.e.,

$$(26) \qquad y = \lambda\eta + \varepsilon,$$

$$(27) \qquad x = \Lambda\xi + \delta$$

where λ is a column vector and Λ is a matrix of factor loadings and ε and δ are the error scores of y and x respectively. The model defined by (25), (26) and (27) contains two kinds of errors namely the error in the equation, i.e., the residual or random disturbance term ζ and the errors in variables, i.e., the measurement errors ε and δ. If one has two or more congeneric or parallel measures of each true score one can estimate both kinds of errors as well as the regression parameters γ. The covariance structure for this model is

$$(28) \qquad \Sigma = \begin{pmatrix} \lambda(\gamma'\Phi\gamma + \psi^2)\lambda' + \Theta_\varepsilon & \\ \Lambda\Phi\gamma'\lambda & \Lambda\Phi\Lambda' + \Theta_\delta \end{pmatrix},$$

where $\Phi = \text{cov}(\xi)$, $\psi^2 = \text{var}(\zeta)$, $\Theta_\varepsilon = \text{cov}(\varepsilon)$, diagonal and $\Theta_\delta = \text{cov}(\delta)$, diagonal.

This model will be illustrated by an example from Rock et al. [1977], in which the regression of role behavior (η) of managers of farm cooperatives on

$\xi_1 = $ knowledge
$\xi_2 = $ value orientation
$\xi_3 = $ role satisfaction
$\xi_4 = $ past training

is estimated. The observed variables are

$y_1 = $ a split-half measure of role behavior
$y_2 = $ a split-half measure of role behavior
$x_{11} = $ a split-half measure of knowledge
$x_{12} = $ a split-half measure of knowledge
$x_{21} = $ a split-half measure of value orientation
$x_{22} = $ a split-half measure of value orientation
$x_{31} = $ a split-half measure of role satisfaction
$x_{32} = $ a split-half measure of role satisfaction
$x_4 = \xi_4 = $ a measure of past training.

For further information about the measures see Warren et al. [1974]. The covariance matrix of all the observed measures is

	y_1	y_2	x_{11}	x_{12}	x_{21}	x_{22}	x_{31}	x_{32}	x_4
y_1	.0271								
y_2	.0172	.0222							
x_{11}	.0219	.0193	.0876						
x_{12}	.0164	.0130	.0317	.0568					
x_{21}	.0284	.0294	.0383	.0151	.1826				
x_{22}	.0217	.0185	.0356	.0230	.0774	.1473			
x_{31}	.0083	.0011	−.0001	.0055	−.0087	−.0069	1137		
x_{32}	.0074	.0015	.0035	.0089	−.0007	−.0088	.0722	.1024	
x_4	.0180	.0194	.0203	.0182	.0563	.0142	−.0056	−.0077	.0946

This can be used to estimate the true regression equation

$$\eta = \gamma_1\xi_1 + \gamma_2\xi_2 + \gamma_3\xi_3 + \gamma_4\xi_4 + \zeta$$

using the following measurement models

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}\eta + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \end{pmatrix}$$

$$\begin{bmatrix} x_{11} \\ x_{12} \\ x_{21} \\ x_{22} \\ x_{31} \\ x_{32} \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1.2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \end{bmatrix} + \begin{bmatrix} \delta_{11} \\ \delta_{12} \\ \delta_{21} \\ \delta_{22} \\ \delta_{31} \\ \delta_{32} \\ 0 \end{bmatrix} .$$

The value 1.2 in the last equation reflects the fact that x_{32} has six items whereas x_{31} has only five.

The overall fit of the model is $\chi^2 = 26.97$ with 22 degrees of freedom, which represents a rather good fit. The ML estimates of the γ's and their standard errors (below) are

$$\hat{\gamma} = (0.350, 0.168, 0.045, 0.071)$$
$$\qquad 0.133 \quad 0.079 \quad 0.054 \quad 0.045 \qquad .$$

These may be compared with the ordinary least squares (OLS) estimates for the regression of y on x_1, x_2, x_3 and x_4, where $y = (1/2)(y_1 + y_2)$, $x_i = (1/2)(x_{i1} + x_{i2})$, $i = 1, 2, 3$. These estimates are

$$\hat{\gamma} = (0.230), 0.120, 0.056, 0.110)$$
$$\qquad 0.052 \quad 0.037 \quad 0.037 \quad 0.038 \qquad .$$

Thus there is considerable bias in the OLS estimates but their standard errors are smaller.

Estimates of the true and error score variances for each observed measure are also obtained. These can be used to compute the reliabilities of the composite measures. The reliability estimates are

$$\begin{array}{cccc} y & x_1 & x_2 & x_3 \\ 0.820 & 0.597 & 0.637 & 0.807 \end{array} .$$

The model defined by (28) can be generalized directly to the case when there are several jointly dependent variables $\mathbf{n}$. The only differences will be that λ and γ' will be replaced by matrices Λ_y and Γ, respectively and ψ^2 by a full symmetric positive definite matrix Ψ [see Jöreskog & Sörbom, Note 2].

7. Path Analysis Models

Path analysis, due to Wright [1934], is a technique sometimes used to assess the direct causal contribution of one variable to another in a non-experimental situation. The problem in general is that of estimating the parameters of a set of linear structural equations representing the cause and effect relationships hypothesized by the investigator. Recently, several models have been studied which involve hypothetical constructs, i.e. latent variables which, while not directly observed, have operational implications for relationships among observable variables [see e.g. Werts & Linn, 1970; Hauser & Goldberger, 1971; Jöreskog & Goldberger, 1975]. Various types of structural equation models in the social sciences were discussed by Jöreskog [1977b] who also considered various statistical problems associated with such models. A computer program for estimation of

structural equation models have been developed by Jöreskog and Sörbom [1978]. In some models, the observed variables appear only as effects (indicators) of the hypothetical constructs, while in others, the observed variables appear as causes or as both causes and effects of latent variables.

Suppose that two variables are used on two occasions, i.e., in a two-wave longitudinal design. Assume that the two variables measure the same latent variable η on two different occasions, i.e., y_1 and y_2 measure η_1 on the first occasion and y_3 and y_4 measure η_2 on the second occasion. The equations defining the measurement relations are

$$(29) \qquad \begin{cases} y_1 = \eta_1 + \epsilon_1 \\ y_2 = \lambda_1 \eta_1 + \epsilon_2 \\ y_3 = \eta_2 + \epsilon_3 \\ y_4 = \lambda_2 \eta_2 + \epsilon_4 \, . \end{cases}$$

The main interest is in the stability of η over time. This can be studied by means of

$$(30) \qquad \eta_2 = \beta\eta_1 + \zeta,$$

the regression of η_2 on η_1. In particular, one is interested in whether β is close to one and ζ is small.

Let Φ be the covariance matrix of (η_1, η_2) and let Θ be the covariance matrix of $(\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4)$. If all the ϵ's are uncorrelated so that Θ is diagonal, the covariance matrix of (y_1, y_2, y_3, y_4) is

$$\Sigma = \begin{bmatrix} \phi_{11} + \theta_{11} & & & \\ \lambda_1\phi_{11} & \lambda_1^2\phi_{11} + \theta_{22} & & \\ \phi_{21} & \lambda_1\phi_{21} & \phi_{22} + \theta_{33} & \\ \lambda_2\phi_{21} & \lambda_1\lambda_2\phi_{21} & \lambda_2\phi_{22} & \lambda_2^2\phi_{22} + \theta_{44} \end{bmatrix} .$$

The matrix Σ has 10 variances and covariances which are functions of 9 parameters. The model has one degree of freedom.

Often when the same variables are used repeatedly there is a tendency for the corresponding errors (the ϵ's) to correlate over time because of memory and other retest effects. Hence there is a need to generalize the above model to allow for correlations between ϵ_1 and ϵ_3 and also between ϵ_2 and ϵ_4. This means that there will be two non-zero covariances θ_{31} and θ_{42} in Θ. The covariance matrix of the observed variables will now be

$$(31) \qquad \Sigma = \begin{bmatrix} \phi_{11} + \theta_{11} & & & \\ \lambda_1\phi_{11} & \lambda_1^2\phi_{11} + \theta_{22} & & \\ \phi_{21} + \theta_{31} & \lambda_1\phi_{21} & \phi_{22} + \theta_{33} & \\ \lambda_2\phi_{21} & \lambda_1\lambda_2\phi_{21} + \theta_{42} & \lambda_2\phi_{22} & \lambda_2^2\phi_{22} + \theta_{44} \end{bmatrix}$$

This Σ has its 10 independent elements expressed in terms of 11 parameters. Hence it is clear that the model is not identified. In fact, none of the 11 parameters are identified without further conditions imposed. The loadings λ_1 and λ_2 may be multiplied by a constant and the ϕ's divided by the same constant. This does not change σ_{21}, σ_{32}, σ_{41} and σ_{43}. The change in the other σ's may be compensated by adjusting the θ's additively. Hence to make the model identified one must fix one λ or one ϕ at a non-zero value or one θ at some arbitrary value. However, the *correlation* between η_1 and η_2 is identified without any restrictions, since

$$\text{corr } (\eta_1, \eta_2) = (\phi_{21}^2/\phi_{11}\phi_{22})^{1/2} = [(\sigma_{32}\sigma_{41})/(\sigma_{21}\sigma_{43})]^{1/2}.$$

This model may therefore be used to estimate this correlation coefficient and to test whether this is one. The maximum likelihood estimate of the correlation coefficient is

$[(s_{32}s_{41})/(s_{21}s_{43})]^{1/2}$. To make further use of the model it is necessary to make some assumption about the nature of the variables. For example, if it can be assumed that the two variables on each occasion are tau-equivalent we can set both λ_1 and λ_2 equal to one. Then the model can be estimated and tested with one degree of freedom. If $\lambda_1 = \lambda_2$, the model is just identified.

While the above model is not identified as it stands it becomes so as soon as there is information about one or more background variables affecting η_1 or η_2 or both. To illustrate this I use an example of a longitudinal study analyzed in more detail by Wheaton et al. [1977]. This study was concerned with the stability over time of attitudes such as alienation and the relation to background variables such as education and occupation. Data on attitude scales were collected from 932 persons in two rural regions in Illinois at three points in time: 1966, 1967 and 1971. The variables used for the present illustration are the *Anomia* subscale and the *Powerlessness* subscale, taken to be indicators of *Alienation*. I use these subscales from 1967 and 1971 only. The background variables are the respondent's education (years of schooling completed) and Duncan's Socioeconomic Index (SEI). These are taken to be indicators of the respondent's socioeconomic status (SES). Let

$$
\begin{aligned}
y_1 &= \text{Anomia 67} & y_3 &= \text{Anomia 71} \\
y_2 &= \text{Powerlessness 67} & y_4 &= \text{Powerlessness 71} \\
x_1 &= \text{Education} & x_2 &= \text{SEI} \\
\zeta &= \text{SES} & \eta_1 &= \text{Alienation 67} \\
& & \eta_2 &= \text{Alienation 71}
\end{aligned}
$$

The model is then specified as

$$
\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \lambda_1 & 0 \\ 0 & 1 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \end{pmatrix},
$$

$$
\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{pmatrix} 1 \\ \lambda_3 \end{pmatrix} \xi + \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix},
$$

$$
\begin{bmatrix} 1 & 0 \\ \beta & 1 \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} = \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix} \xi + \begin{bmatrix} \zeta_1 \\ \xi_2 \end{bmatrix}.
$$

It is assumed that ζ_1 and ζ_2 are uncorrelated and that the scales for η_1, η_2 and ξ have been chosen to be the same as for y_1, y_3 and x_1, respectively.

Let $\phi = \text{var}(\xi)$ and $\psi_i = \text{var}(\zeta_i)$, $i = 1, 2$ and let $\boldsymbol{\Phi}$ be the covariance matrix of (η_1, η_2, ξ). It is obvious that there is a one-to-one correspondence between the six ϕ's in $\boldsymbol{\Phi}$ and $(\phi, \beta, \gamma_1, \gamma_2, \psi_1, \psi_2)$. In terms of $\boldsymbol{\Phi}$, the covariance matrix of $(y_1, y_2, y_3, y_4, x_1, x_2)$ is

$$
\boldsymbol{\Sigma} = \begin{bmatrix}
\phi_{11} + \theta_{11} & & & & & \\
\lambda_1\phi_{11} & \lambda_1^2\phi_{11} + \theta_{22} & & & & \\
\phi_{21} + \theta_{31} & \lambda_1\phi_{21} & \phi_{22} + \theta_{33} & & & \\
\lambda_2\phi_{21} & \lambda_1\lambda_2\phi_{21} + \theta_{42} & \lambda_2\phi_{22} & \lambda_2^2\phi_{22} + \theta_{44} & & \\
\phi_{31} & \lambda_2\phi_{31} & \phi_{32} & \lambda_2\phi_{32} & \phi_{33} + \theta_{55} & \\
\lambda_3\phi_{31} & \lambda_1\lambda_3\phi_{31} & \lambda_3\phi_{32} & \lambda_2\lambda_3\phi_{32} & \lambda_3\phi_{33} & \lambda_3^2\phi_{33} + \theta_{66}
\end{bmatrix}
$$

The upper left 4×4 part of the matrix is the same as (31). It is clear that the two last rows of Σ determine λ_1, λ_2, λ_3, ϕ_{31}, ϕ_{32}, ϕ_{33}, θ_{55} and θ_{66}. With λ_1 and λ_2 determined, the other parameters are determined by the upper left part. Altogether we have 17 parameters to estimate so that there are 4 degrees of freedom.

The sample covariance matrix of the six observed variables is ($N = 932$):

$$
\mathbf{S} = \begin{array}{c} y_1 \\ y_2 \\ y_3 \\ y_4 \\ x_1 \\ x_2 \end{array}
\begin{bmatrix}
y_1 & y_2 & y_3 & y_4 & x_1 & x_2 \\
11.834 & & & & & \\
6.947 & 9.364 & & & & \\
6.819 & 5.091 & 12.532 & & & \\
4.783 & 5.028 & 7.495 & 9.986 & & \\
-3.839 & -3.889 & -3.841 & -3.625 & 9.610 & \\
-21.899 & -18.831 & -21.748 & -18.775 & 35.522 & 450.288
\end{bmatrix} .
$$

These data have been analyzed under two models. In Model A, $\theta_{31} = \theta_{42} = 0$; in Model B, θ_{31} and θ_{42} are free.

The maximum likelihood estimates of the parameters with standard errors in parenthesis are given in Table 6. The stability of alienation over time is reflected in the parameter β. The influence of SES on Alienation at the two occasions is significant in Model A. The coefficient for 1967, γ_1, -0.614 with a standard error of 0.056 and for 1971, γ_2, it is -0.174 with a standard error equal to 0.054. The negative sign of the γ-coefficients indicates that for high socioeconomic status the alientation is low and vice versa. However, the overall fit of the Model A is not acceptable: χ^2 with six degrees of freedom equals 71.544. Model B is intuitively more plausible. As can be seen from Table 6, the inclusion of θ_{31} and θ_{42} results in a model with an acceptable overall fit. A test of the hypothesis that both θ_{31} and θ_{42} are zero yields $\chi^2 = 66.774$ with 2 degrees of freedom so that this hypothesis must be rejected.

Other similar models for longitudinal data with two or more occasions have been studied by Jöreskog and Sörbom [1977] and Jöreskog [1978].

8. Simplex and Circumplex Models

Simplex Models

Since the fundamental paper of Guttman [1954] on the simplex structure for correlations between ordered tests, many investigators have found data displaying the typical simplex structure. Guttman gave several examples of this structure. His Table 5 is reproduced here as Table 7. In this example all the tests involve verbal ability and are ordered according to increasing complexity.

The typical property in a simplex correlation structure, such as that in Table 7, is that the correlations decrease as one moves away from the main diagonal. Such data will not usually fit a factor analysis model with one common factor for the following reasons. Let the tests be ordered so that the factor loadings decrease. Then if the factor model holds, the correlations in the first row decrease as one moves away from the diagonal, but the correlations in the last row *increase* as one moves away from the diagonal. Also the correlations just below the diagonal *decrease markedly* as one moves down. These features do not hold in Table 7.

Jöreskog [1970b] considered several statistical models for such simplex structures. Following Anderson [1960] he formulated these models in terms of the well-known Wiener and Markov stochastic processes. A distinction was made between a perfect simplex and a quasi simplex. A perfect simplex is reasonable only if the measurement

TABLE 6

Maximum Likelihood Estimates for Models A and B

The standard errors of the estimates are given within parenthesis.

Parameter	Model A	Model B
λ_1	0.889 (.041)	0.979 (.062)
λ_2	0.849 (.040)	0.922 (.060)
λ_3	5.329 (.430)	5.221 (.422)
β	0.705 (.054)	0.607 (.051)
γ_1	−0.614 (.056)	−0.575 (.056)
γ_2	−0.174 (.054)	−0.227 (.052)
ψ_1	5.307 (.473)	4.846 (.468)
ψ_2	3.742 (.388)	4.089 (.405)
ϕ	6.666 (.641)	6.803 (.650)
θ_{11}	4.015 (.343)	4.735 (.454)
θ_{22}	3.192 (.271)	2.566 (.404)
θ_{33}	3.701 (.373)	4.403 (.516)
θ_{44}	3.625 (.292)	3.074 (.435)
θ_{31}	− − −	1.624 (.314)
θ_{42}	− − −	0.339 (.261)
θ_{55}	2.944 (.500)	2.807 (.508)
θ_{66}	260.982 (18.242)	264.809 (18.154)
χ^2	71.470	4.730
d.f.	6	4

TABLE 7

Intercorrelations of Six Verbal-Ability Tests

for 1046 Bucknell College Sophomores

Test	Spelling	Punctuation	Grammar	Vocabulary	Literature	Foreign Literature
	A	C	B	D	E	H
A	--	.621	.564	.476	.394	.389
C	.621	--	.742	.503	.461	.411
B	.564	.742	--	.577	.472	.429
D	.476	.503	.577	--	.688	.548
E	.394	.461	.472	.688	--	.639
H	.389	.411	.429	.548	.639	--
Total	2.444	2.738	2.784	2.792	2.654	2.416

errors in the test scores are negligible. A quasi simplex, on the other hand, allows for sizeable errors of measurement. Simplex models occur naturally when the same or contentwise similar measurements are repeated on the same individuals over time. In the following we discuss all models in terms of four occasions, the generalization to an arbitrary number of occasions will be obvious at all stages.

The unit of measurement in the true variables η_i may be chosen to be the same as in the observed variables y_i, $i = 1, 2, 3, 4$. The equations defining the model are then, taking all variables as deviations from their means:

$$(32) \qquad y_i = n_i + \epsilon_i, \qquad i = 1, 2, 3, 4,$$

$$(33) \qquad \eta_i = \beta_i \eta_{i-1} + \zeta_i, \qquad i = 2, 3, 4,$$

where the ϵ_i are uncorrelated among themselves and uncorrelated with all the η_i and where ζ_{i+1} is uncorrelated with η_i, $i = 1, 2, 3$. The parameters of the model are $\phi_i = \text{var}(\eta_i)$, $\theta_i = \text{var}(\epsilon_i)$, $i = 1, 2, 3, 4$ and $\beta_2, \beta_3, \beta_4$. The residual variance var (ζ_{i+1}) is a function of ϕ_{i+1}, ϕ_i and β_{i+1}, namely var $(\zeta_{i+1}) = \phi_{i+1} - \beta_{i+1}^2 \phi_i$, $i = 1, 2, 3$. The covariance matrix of y_1, y_2, y_3 and y_4 is

$$(34) \qquad \Sigma = \begin{bmatrix} \phi_1 + \theta_1 & & & \\ \beta_2 \phi_1 & \phi_2 + \theta_2 & & \\ \beta_2 \beta_3 \phi_1 & \beta_3 \phi_2 & \phi_3 + \theta_3 & \\ \beta_2 \beta_3 \beta_4 \phi_1 & \beta_3 \beta_4 \phi_2 & \beta_4 \phi_3 & \phi_4 + \theta_4 \end{bmatrix}.$$

It is seen from (34) that although the product $\beta_2 \phi_1 = \sigma_{21}$ is identified, β_2 and ϕ_1 are not separately identified. The product $\beta_2 \phi_1$ is involved in the off-diagonal elements in the first column (and row) only. One can multiply β_2 by a constant and divide ϕ_1 by the same constant without changing the product. The change induced by ϕ_1 in σ_{11} can be absorbed in θ_1 in such a way that σ_{11} remains unchanged. Hence $\theta_1 = \text{var}(\epsilon_1)$ is not identified. For η_2

and η_3 we have

$$\phi_2 = \frac{\sigma_{32}\sigma_{21}}{\sigma_{31}},$$

$$\phi_3 = \frac{\sigma_{43}\sigma_{32}}{\sigma_{42}},$$

so that ϕ_2 and ϕ_3, and hence also θ_2 and θ_3, are identified. With ϕ_2 and ϕ_3 identified, β_3 and β_4 are identified by σ_{32} and σ_{43}. The middle coefficient β_3 is overidentified since

$$\beta_3\phi_2 = \frac{\sigma_{31}\sigma_{42}}{\sigma_{41}} = \sigma_{32}.$$

Since both ϕ_4 and θ_4 are involved in σ_{44} only, they are not identified but their sum σ_{44} is identified.

This analysis of the identification problem shows that for the "inner" variables y_2 and y_3, the parameters ϕ_2, ϕ_3, θ_{22}, θ_{33} and β_3 are identified, whereas there is an indeterminacy associated with each of the "outer" variables y_1 and y_4. To eliminate these indeterminacies, one of the parameters ϕ_1, θ_1 and β_2 must be specified and one of the parameters ϕ_4 and θ_4 must also be specified. Perhaps the most natural way of eliminating the indeterminacies is to set $\theta_1 = \theta_2$ and $\theta_4 = \theta_3$. Hence there are only nine independent parameters and the model has one degree of freedom. In the general case of $p \geq 4$ occasions there will be $3p - 3$ free parameters and the degrees of freedom are $(1/2)p(p + 1) - (3p - 3)$.

To illustrate a simplex model, I will use data published by Humphreys [1968] and analyzed by Werts et al. [1978]. The variables include eight semesters of gradepoint averages, high school rank and a composite score on the American College Testing tests for approximately 1600 undergraduate students at the University of Illinois. The correlation matrix is given in Table 8.

Using first only $y_1 - y_8$, I estimate a quasi-Markov simplex and estimate the correlations between the true academic achievements η_2, η_3, $\cdots$, η_8. These correlations are:

	η_2	η_3	η_4	η_5	η_6	η_7
η_2	1.000					
η_3	0.838	1.000				
η_4	0.812	0.969	1.000			
η_5	0.724	0.865	0.892	1.000		
η_6	0.677	0.809	0.834	0.935	1.000	
η_7	0.619	0.740	0.763	0.855	0.914	1.000

Here every correlation ρ_{ij} with $|i - j| > 1$ is the product of correlations just below the diagonal. For example, $\rho(\eta_5, \eta_2) = 0.838 \cdot 0.969 \cdot 0.892 = 0.724$. These correlations form a perfect Markov simplex. The goodness of fit test of the model gives $\chi^2 = 23.91$ with 15 degrees of freedom. This represents a reasonably good fit considering the large sample size. The reliabilities of the semester grades y_2, y_3, $\cdots$, y_7 can also be obtained directly from the solution in which the η's are standardized. The reliabilities are

y_2	y_3	y_4	y_5	y_6	y_7
0.569	0.575	0.562	0.584	0.581	0.608.

A test of the hypothesis that these are equal gives $\chi^2 = 2.17$ with 5 degrees of freedom, so that this hypothesis is not rejected by the data despite the large sample size.

In this example the correlations $\rho(\eta_1, \eta_j)$, $j \neq 1$ and $\rho(\eta_i, \eta_8)$, $i \neq 8$ and the reliabilities of y_1 and y_8 are not identified. However, in view of the above test of equality of reliabilities

TABLE 8

Correlations among Grade Point Averages, High
School Rank and An Aptitude Test

	y_0	y_0'	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8
y_0	1.000									
y_0'	.393	1.000								
y_1	.387	.375	1.000							
y_2	.341	.298	.556	1.000						
y_3	.278	.237	.456	.490	1.000					
y_4	.270	.255	.439	.445	.562	1.000				
y_5	.240	.238	.415	.418	.496	.512	1.000			
y_6	.256	.252	.399	.383	.456	.469	.551	1.000		
y_7	.240	.219	.387	.364	.445	.442	.500	.544	1.000	
y_8	.222	.173	.342	.339	.354	.416	.453	.482	.541	1.000

Note: y_0 is high school rank, y_0' ACT composite score, and
y_1 through y_8 are eight semesters grade-point averages.

it seems reasonable to assume that all reliabilities or equivalently all error variances in the standardized solution are equal for y_1 through y_8. This assumption makes it possible to estimate the intercorrelations among all the η's.

Assuming that x_0 and x_0' are indicators of precollege academic achievement η_0 which influences the true academic achievement in the first semester η_1, one can estimate again the quasi-Markov simplex and show how this use of x_0 and x_0' helps identify the parameters of the model.

The only parameter which is now not identified is θ_8, the error variance in y_8. This gives a $\chi^2 = 36.92$ with 28 degrees of freedom. If we assume that the reliabilities of all the semester grades are equal, all parameters are identified and the goodness of fit becomes 45.22 with 34 degrees of freedom. The difference equaling 8.30 with 6 degrees of freedom provides another test of equality of the reliabilities. Finally a test of the hypothesis that the whole process is stationary, i.e., that

$$\beta_2 = \beta_3 = \cdots = \beta_8$$

$$\theta_2 = \theta_3 = \cdots = \theta_8$$

gives $\chi^2 = 12.82$ with 11 degrees of freedom so that this hypothesis cannot be rejected. There is good evidence that the whole Markov process is stable over time.

Circumplex Models

Simplex models, as considered in the previous subsection, are models for tests that may be conceived of as having a linear ordering. The circumplex is another model

considered by Guttman [1954] and this yields a circular ordering instead of a linear. The circular order has no beginning and no end but there is still a law of neighboring that holds.

The circumplex model suggested by Guttman [1954] is a circular moving average process. Let $\zeta_1, \zeta_2, \cdots, \zeta_p$ be uncorrelated random latent variables. Then the perfect circumplex of order m with p variables is defined by

$$x_i = \zeta_i + \zeta_{i+1} + \cdots + \zeta_{i+m-1},$$

where $x_{p+i} = x_i$. In matrix form we may write this as $\mathbf{x} = \mathbf{Cf}$, where $\mathbf{C}$ is a matrix of order $p \times p$ with zeros and ones. In the case of $p = 6$ and $m = 3$,

$$\mathbf{C} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Let $\phi_1, \phi_2, \cdots, \phi_p$ be the variances of $\zeta_1, \zeta_2, \cdots, \zeta_p$ respectively. Then the covariance matrix of $\mathbf{x}$ is

$$\mathbf{\Sigma} = \mathbf{CD\,C'}$$

where $\mathbf{D} = \text{diag}\,(\phi_1, \phi_2, \cdots, \phi_p)$. The variance of x_i is

$$\sum_{k=i}^{i+m-1} \phi_k,$$

the covariance between x_i and x_j, for $i < j$, is

$$\sum_{k=j}^{i+m-1} \phi_k \quad \text{for} \quad j = i+1, i+2, \cdots, i+m-1$$

and 0 otherwise and the correlation between x_i and x_j, $i < j$, is

$$\rho_{ij} = \frac{\displaystyle\sum_{k=j}^{i+m-1} \phi_k}{\left[\left(\displaystyle\sum_{k=i}^{i+m-1} \phi_k\right)\left(\displaystyle\sum_{k=j}^{j+m-1} \phi_k\right)\right]^{1/2}}.$$

Here ϕ_{p+k} should be interpreted as ϕ_k. If all the ϕ's are equal, i.e., if the points on the circle are equidistant, the correlations form a circular symmetric pattern as follows:

$$(35) \qquad \mathbf{P} = \begin{bmatrix} 1 & & & & & \\ \rho_1 & 1 & & & & \\ \rho_2 & \rho_1 & 1 & & & \\ \rho_3 & \rho_2 & \rho_1 & 1 & & \\ \rho_2 & \rho_3 & \rho_2 & \rho_1 & 1 & \\ \rho_1 & \rho_2 & \rho_3 & \rho_2 & \rho_1 & 1 \end{bmatrix}.$$

Even with different ϕ's the correlations ρ_{ij}, for a given i, decrease as j increases beyond i, reach a minimum, and then increase as j approaches $p + i$. It is convenient to think of $1, 2, \cdots, p$ as points on a circle. Then for adjacent points the correlation tends to be high, and for points far apart the correlation tends to be low or 0. If $m \le p/2$, then $\rho_{ij} = 0$ if $j - i > m$ (modulus p). Since zero correlations are not to be expected in practice, we assume that m is chosen to be greater than $p/2$.

TABLE 9

Intercorrelations Among Tests of Six Different Kinds of

Abilities for 710 Chicago Schoolchildren

Test	1	2	3	4	5	6
1. Association	1.000	0.446	0.321	0.213	0.234	0.442
2. Incomplete Words	0.446	1.000	0.388	0.313	0.208	0.330
3. Multiplication	0.321	0.388	1.000	0.396	0.325	0.328
4. Dot Patterns	0.213	0.313	0.396	1.000	0.352	0.247
5. ABC	0.234	0.208	0.325	0.352	1.000	0.347
6. Directions	0.442	0.330	0.328	0.247	0.347	1.000

Guttman gives several examples of sets of tests showing this property of the correlations. His Table 19 is reproduced here as Table 9. It will be noted that in each row the correlations fall away from the main diagonal, reach a minimum, and then increase. Guttman argues that the tests form a circular order. In the example the Association test is about equally related to the Incomplete Words test and the Directions test. It is this circular ordering that is the significant feature of the correlation matrix.

The perfect circumplex model may be too restrictive. First, it cannot account for random measurement error in the test scores and second, since the model is not scale free, it cannot be used to analyze correlations such as those in Table 9. However, these difficulties are easily remedied by considering the following quasi circumplex

$$x = D_\alpha C \zeta + e$$

with covariance matrix

(36) $$\Sigma = D_\alpha C D_\phi C' D_\alpha + \Theta^2 ,$$

where e is the vector of error scores with variances in the diagonal matrix Θ^2 and D_α is a diagonal matrix of scale factors. One element in D_α or D_ϕ must be fixed at unity. It seems most natural to fix $\phi_1 = 1$.

To illustrate the testing of a circumplex model, I use the data in Table 9. I first test circular correlation pattern with the model

$$\Sigma = D_\sigma P D_\sigma ,$$

where P is given by (35) and where D_σ is a diagonal matrix of standard deviations. Note that D_σ must be estimated even though S is a correlation matrix. The model gives $\chi^2 = 27.05$ with 12 degrees of freedom. The more general model (36) with $m = 4$ gives $\chi^2 = 16.47$ with 4 degrees of freedom. It appears that these data do not fit a circumplex model well.

REFERENCE NOTES

1. Gruveaus, G. & Jöreskog, K. G. *A computer program for minimizing a function of several variables* (Res. Bull. 70-14). Princeton, N. J.: Educational Testing Service, 1970.
2. Jöreskog, K. G. & Sörbom, D. *Some regression models useful in the measurement of change* (Res. Rep. 74-3). Uppsala, Sweden: Department of Statistics, University of Uppsala, 1974.

REFERENCES

Aitken, A. C. On least squares and the linear combination of observations. *Proceedings of the Royal Society of Edinburgh,* 1934-35, *55,* 42-28.

Anderson, T. W. Some stochastic process models for intelligence test scores. In K. J. Arrow, S. Karlin & P. Suppes (Eds.), *Mathematical methods in the social sciences, 1959.* Stanford, Calif.: Stanford University Press, 1960, 205–220.

Anderson, T. W. The use of factor analysis in the statistical analysis of multiple time series. *Psychometrika,* 1963, *28,* 1–25.

Anderson, T. W. Statistical inference for covariance matrices with linear structure. In P. R. Krishnaiah (Ed.), *Multivariate analysis—II.* New York: Academic Press, 1969, 55–66.

Archer, C. O. & Jennrich, R. I. Standard errors for rotated factor loadings. *Psychometrika,* 1973, *38,* 581–592.

Bock, R. D. Components of variance analysis as a structural and discriminal analysis for psychological tests. *British Journal of Statistical Psychology,* 1960, *13,* 151–163.

Bock, R. D. & Bargmann, R. E. Analysis of covariance structures. *Psychometrika,* 1966, *31,* 507–534.

Browne, M. W. Generalized least squares estimators in the analysis of covariance structures. *South African Statistical Journal,* 1974, *8,* 1–24. (Reprinted in D. J. Aigner & A. S. Goldberger (Eds.), *Latent variables in socio-economic models.* North Holland Publishing Co., 1977, 203–226.)

Browne, M. W. The analysis of patterned correlation matrices by generalized least squares. *British Journal of Mathematical and Statistical Psychology,* 1977, *30,* 113–124.

Campbell, D. T. & Fiske, D. W. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin,* 1959, *56,* 81–105.

Fletcher, R. & Powell, M. J. D. A rapidly convergent descent method for minimization. *The Computer Journal,* 1963, *6,* 163–168.

French, J. W. The description of aptitude and achievement tests in terms of rotated factors. *Psychometric Monographs,* 5, 1951.

Graybill, F. A. *An introduction to linear statistical models* (Vol. 1). New York: McGraw-Hill, 1961.

Guilford, J. P. The structure of intellect. *Psychological Bulletin,* 1956, *53,* 267–293.

Guttman, L. A new approach to factor analysis: The radex. In P. F. Lazarsfeld (Ed.), *Mathematical thinking in the social sciences.* New York: Columbia University Press, 1954, 258–348.

Harman, H. H. *Modern factor analysis.* Chicago: The University of Chicago Press, 1967.

Hauser, R. M. & Goldberger, A. S. The treatment of unobservable variables in path analysis. In H. L. Costner (Ed.), *Sociological Methodology 1971.* London: Jossey-Bass, 1971, 81–117.

Hendrickson, A. E. & White, P.O. Promax: A quick method for rotation to oblique simple structure. *British Journal of Statistical Psychology,* 1964, *17,* 65–70.

Humphreys, L. G. The fleeting nature of college academic success. *Journal of Educational Psychology,* 1968, *59,* 375–380.

Jennrich, R. I. Standard errors for obliquely rotated factor loadings. *Psychometrika,* 1973, *38,* 593–604.

Jöreskog, K. G. Some contributions to maximum likelihood factor analysis. *Psychometrika,* 1967, *32,* 443–482.

Jöreskog, K. G. A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika,* 1969, *34,* 183–202.

Jöreskog, K. G. A general method for analysis of covariance structures. *Biometrika,* 1970, *57,* 239–251. (a)

Jöreskog, K. G. Estimation and testing of simplex models. *British Journal of Mathematical and Statistical Psychology,* 1970, *23,* 121–145. (b)

Jöreskog, K. G. Analysis of covariance structures. In P. R. Krishnaiah (Ed.), *Multivariate analysis—III.* New York: Academic Press, Inc., 1973, 263–285.

Jöreskog, K. G. Analyzing psychological data by structural analysis of covariance matrices. In R. C. Atkinson, D. H. Krantz, R. D. Luce & P. Suppes (Eds.), *Contemporary developments in mathematical psychology—Volume II.* San Francisco: W. H. Freeman & Co., 1974, 1–56.

Jöreskog, K. G. Factor analysis by least squares and maximum likelihood methods. K. Enslein, A. Ralston & M. S. Wilf (Eds.), *Statistical methods for digital computers.* New York: Wiley, 1977, 125–165. (a)

Jöreskog, K. G. Structural equation models in the social sciences: Specification, estimation and testing. In P. R. Krishnaiah (Ed.), *Applications of statistics.* Amsterdam: North Holland Publishing Co., 1977, 265–286. (b)

Jöreskog, K. G. Statistical estimation of structural models in longitudinal-developmental investigations. In J. R. Nesselroade & P. B. Baltes (Eds.), *Longitudinal research in the behavioral sciences: Design and analysis.* New York: Academic Press, 1978, in press.

Jöreskog, K. G. & Goldberger, A. S. Factor analysis by generalized least squares. *Psychometrika,* 1972, *37,* 243–260.

Jöreskog, K. G. & Goldberger, A. S. Estimation of a model with multiple indicators and multiple causes of a single latent variable. *Journal of the American Statistical Association,* 1975, *10,* 631–639.

Jöreskog, K. G. & Sörbom, D. Statistical models and methods for analysis of longitudinal data. In D. J. Aigner & A. S. Goldberger (Eds.), *Latent variables in socio-economic models.* Amsterdam: North Holland Publishing Co., 1977, 235–285.

Jöreskog, K. G. & Sörbom, D. *LISREL IV—A general computer program for estimation of a linear structural equation system by maximum likelihood methods.* Chicago: National Educational Resources, 1978.

Jöreskog, K. G. & Sörbom, D. *EFAP II-Exploratory factor analysis program.* Chicago: National Educational Resources, 1979, in press.

Kaiser, H. F. The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 1958, *23*, 187–200.

Kristof, W. On the theory of a set of tests which differ only in length. *Psychometrika*, 1971, *36*, 207–225.

Lawley, D. N. & Maxwell, A. E. *Factor analysis as a statistical method* (2nd ed.). New York: American Elsevier, 1971.

Lord, F. M. A study of speed factors in tests and academic grades. *Psychometrika*, 1956, *21*, 31–50.

Lord, F. M. & Novick, M. E. *Statistical theories of mental test scores.* Reading: Addison-Wesley Publishing Co., 1968.

Miller, D. M. & Lutz, M. V. Item design for an inventory of teaching practices and learning situations. *Journal of Educational Measurement*, 1966, *3*, 53–61.

McDonald, R. P. Testing pattern hypotheses for covariance matrices. *Psychometrika*, 1974, *39*, 189–20.

McDonald, R. P. Testing pattern hypotheses for correlation matrices. *Psychometrika*, 1975, *40*, 253–255.

Mukherjee, B. N. Likelihood ratio tests of statistical hypotheses associated with patterned covariance matrices in psychology. *British Journal of Mathematical and Statistical Psychology*, 1970, *23*, 120.

Rao, C. R. *Linear statistical inference and its applications* (2nd ed.). New York: Wiley, 1973.

Rock, D. A., Werts, C. E., Linn, R. L. & Jöreskog, K. G. A maximum likelihood solution to the errors in variables and errors in equations model. *Journal of Multivariate Behavioral Research*, 1977, *12*, 187–197.

Thurstone, L. L. Primary mental abilities. *Psychometric Monographs*, *1*, 1938.

Thurstone, L. L. *Multiple factor analysis.* Chicago: University of Chicago Press, 1947.

Warren, R. D., White, J. K. & Fuller, W. A. An errors in variables analysis of managerial role performance. *Journal of the American Statistical Association*, 1974, *69*, 886–893.

Werts, C. E. & Linn, R. L. Path analysis: Psychological examples. *Psychological Bulletin*, 1970, *67*, 193–212.

Werts, C. E., Linn, R. L. & Jöreskog, K. G. Reliability of college grades from longitudinal data. *Educational and Psychological Measurement*, 1978, *38*, 89–95.

Wheaton, B., Muthén, B., Alwin, D. & Summers, G. Assessing reliability and stability in panel models. In D. R. Heise (Ed.), *Sociological Methodology 1977.* San Francisco: Jossey-Bass, 1977.

Wiley, D. E., Schmidt, W. H. & Bramble, W. J. Studies of a class of covariance structure models. *Journal of the American Statistical Association*, 1973, *68*, 317–323.

Wright, S. The method of path coefficients. *The Annals of Mathematical Statistics*, 1934, *5*, 161–215.