1216

THE USE OF FACTOR ANALYSIS IN RESEARCH

by

65

JAMES IVEY CLOGSTON

B. S., Friends University, 1957

A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Statistics

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1964

Approved by:

Gary F. Krause
Major Professor

TABLE OF CONTENTS

INTRODUCTION

Factor Analysis is a branch of statistical science. It is often mistakenly considered as psychological theory since the factor analysis technique was originally developed and extensively used in the field of psychology. The method came into being originally to provide mathematical models for the explanation of psychological theories of human ability and behavior.

The mathematical techniques inherent in factor analysis are not limited to psychological applications. The principal concern of factor analysis is the resolution of a set of variables linearly in terms of a small number of categories called "factors". This resolution can be accomplished by the analysis of the correlations among the variables. A satisfactory solution will yield factors which convey all the essential information of the original set of variables. The chief aim of factor analysis is to attain economy of description which is sometimes called scientific parsimony.

The factor analysis technique is not new. Charles Spearman is generally given credit for the birth of factor analysis. His paper "General Intelligence, Objectively Determined and Measured" was published in the American Journal of Psychology in 1904. This early work was the beginning of the development of the "Two - factor Theory" although the term factor was not used explicitly at that time. From a statistical point of view an important contribution to factor analysis was made by Karl Pearson who published a paper in 1901 on "The Method of Principal Axes".

From 1904 to 1925 a considerable amount of work on psychological theories and mathematical foundations was completed. In the 30's it became evident that Spearman's Two - factor Theory was not always adequate to describe a battery of psychological tests.

Garnett (1919) explored the possibility of extracting several factors directly from a matrix of correlations among tests and the concept of multiple - factor analysis resulted. In addition to Garnett; Thurstone (1935), Hotelling (1933), Holzinger (1934, 35, 36) and Thomson (1936) made important contributions to the development of multiple factor analysis in the 1930's.

During World War II with large scale testing, classification, and assignment problems research psychologists employed factor analysis widely throughout the military services. In recent years factor analysis has been applied to fields other than that of psychology. Factor analysis has been used in such fields as sociology, meterology, political science, geography, economics, physiology and medicine.

The applications of factor analysis indicated above are concerned primarily with classification and verification of scientific hypotheses in the particular field of investigation. A different application of factor analysis is to supplement and simplify conventional statistical techniques and computations. Typical of this kind of application is the use of factor analysis in expediting the computation of multiple regression statistics.

This paper will be devoted primarily to presenting some of the underlying theory and application of the factor analysis technique.

THEORETICAL ASPECTS

The Model

Since factor analysis involves the correlation matrix among several different variables it is necessary to develop the notation to be used for this report. A measurement from the jth variable and the ith individual will be X_{ji} where $j = 1, \ldots, n$ and $i = 1, \ldots, N$ and the mean

$$\bar{X}_j = \sum_{i=1}^{N} X_{ji}/N \; . \qquad (1)$$

The deviate from the mean will be

$$x_{ji} = X_{ji} - \bar{X}_j \qquad (2)$$

consequently the variance of the variable X_{ji} is

$$\sigma_j^2 = \sum_{i=1}^{N} {x_{ji}}^2 / N \; . \qquad (3)$$

The standardized value of the variable j for the ith individual will be

$$z_{ji} = x_{ji} / \sigma_j \; . \qquad (4)$$

The set of all values z_{ji} $(i = 1, \ldots, N)$ is called a statistical variable z_j in standard form. The variance of this variable is unity.

The product moment-correlation between any two standardized variables say j and k is

$$r_{jk} = \sum_{i=1}^{N} z_{ji}\, z_{ki} \;/ N. \qquad (5)$$

The intercorrelations among all the variables of a study constitute the basic data for factor analysis.

It is the object of the factor analysis to represent variable z_j in terms of several underlying factors. The simplest mathematical model for describing a variable in terms of several others is the linear model. There are two types of factors that will be distinguished.

(a) Common Factors - involved in more than one variable of a set. These factors will be denoted $F_1, F_2, \ldots, F_m$ where $m \leq n$ is the total number of common factors.

(b) Unique Factors - involved in a single variable of a set. $U_1, U_2, \ldots, U_n$ will denote the unique factors.

The linear expression for any variable z_j ($j = 1, 2, \ldots, n$) may be written as

$$z_j = a_{j1}F_1 + a_{j2}F_2 + \ldots + a_{jm}F_m + a_jU_j \ . \qquad (6)$$

This equation may be rewritten explicitly for the ith individual as

$$z_{ji} = a_{j1}F_{1i} + a_{j2}F_{2i} + \ldots + a_{jm}F_{mi} + a_jU_{ji} \ . \qquad (7)$$

By squaring, summing over the N values and dividing by N

$$\sum_{i=1}^{N} z_{ji}^2 / N = a_{j1}^2 \sum_{i=1}^{N} F_{1i}^2 / N + a_{j2}^2 \sum_{i=1}^{N} F_{2i} / N + \ldots$$

$$+ a_{jm}^2 \sum_{i=1}^{N} F_{mi}^2 / N + a_j^2 \sum_{i=1}^{N} U_{ji}^2 / N + \qquad (8)$$

$$2(a_{j1} a_{j2} \sum_{i=1}^{N} F_{1i}F_{2i} / N + \ldots + a_{jm} a_j F_{mi}U_{ji} / N)$$

$$= 1/N \sum_{i=1}^{N} \left[\frac{x_{ji}}{\sigma_j}\right]^2 = \frac{1}{N\sigma_j^2} \sum_{i=1}^{N} (x_{ji})^2 \quad ,$$

$$\text{but } E \left[\frac{1}{N\sigma_j^2} \sum_{i=1}^{N} (x_{ji})^2\right] = \frac{1}{N\sigma_j^2} \cdot N\sigma_j^2 = 1 = \sigma_j^2$$

The variance of any standardized variable z_j is unity.

Variance Components

Since the variance of a variable in standard form is equal to unity, and all variables (including the factors) are assumed to be in standard form for any sample, equation (8) can be rewritten,

$$1 = \sigma_j^2 = a_{j1}^2 + a_{j2}^2 + \cdots + a_{jm}^2 + a_j^2 + a(a_{j1}a_{j2}\, r_{F_1F_2} + \cdots$$

$$+ a_{jm}a_j r_{F_mU_j}) \quad . \tag{9}$$

If the factors are uncorrelated this can be rewritten:

$$1 = \sigma_j^2 = a_{j1}^2 + a_{j2}^2 + \cdots + a_{jm}^2 + a_j^2 \quad . \tag{10}$$

The terms on the right hand side of this equation represent the portions of the unit variance of z_j ascribable to the respective factors. For example a_{j2}^2 is the contribution of the factor F_2 to the variance of z_j. The total contribution of a factor F_p to the variances of all the variables is defined to be $V_p = \sum_{j=1}^{n} a_{jp}^2$ $(p = 1, 2, \ldots, m)$. From the composition of the total unit variance as expressed by (10) two important concepts in factor analysis follow; (a) the communality of a variable z, which is given by the sum of squares of the common factor coefficients,

$$h_j^2 = a_{j1}^2 + a_{j2}^2 + \cdots + a_{jm}^2 \quad (j = 1, 2, \ldots, n) \tag{11}$$

and (b) the contribution of the uniqueness factor. This uniqueness factor can be broken down into two portions. $a_j^2 = b_j^2 + c_j^2$ where b_j^2 is called the specificity of the variable and c_j^2 is called the error variance of the variable. Therefore the total unit variance can be expressed as

$$1 = h_j^2 + a_j^2 = h_j^2 + b_j^2 + c_j^2 \ . \tag{12}$$

We obtain the communality and the uniqueness a_j^2 of each variable in a set by factorial methods. The further splitting of the uniqueness factor is independent of the factorial solution, however the reliability of the variable can generally be found by experimental methods and the error variance is given by the equation

$$c_j^2 = 1 - r_{jj} \tag{13}$$

where r_{jj} is the reliability, and

$$b_j^2 = a_j^2 - c_j^2 \tag{14}$$

where a_j^2 is known from the factor solution.

It follows that

$$r_{jj} = h_j^2 + b_j^2 \tag{15}$$

and

$$h_j^2 = r_{jj} - b_j^2 \quad . \tag{16}$$

In other words the communality is always less than or equal to the reliability of the variable. It becomes equal only when the specificity vanishes.

Employing The foregoing model for the expression of a variable z_j in terms of factors, the components of variance are given by:

Total Variance	$(1) = h_j^2 + b_j^2 + c_j^2$	$= h_j^2 + a_j^2$
Reliability	$(r_{jj}) = h_j^2 + b_j^2$	$= 1 - c_j^2$
Communality	$(h_j^2) = h_j^2$	$= 1 - a_j^2$
Uniqueness	$(a_j^2) = b_j^2 + c_j^2$	$= 1 - h_j^2$
Specificity	$(b_j^2) = b_j^2$	$= a_j^2 - c_j^2$
Error Variance	$(c_j^2) = c_j^2$	$= 1 - r_{jj}$

Objectives

Having described the composition of variables in terms of factors, it is now possible to outline the objectives of a factor analysis of a set of data. For a set of n variables, the linear model (6) may be rewritten in expanded form as follows:

$$\begin{array}{l} z_1 = a_{11}F_1 + a_{12} F_2 + \cdots + a_{1m} F_m + a_1 U_1 \\ z_2 = a_{21}F_1 + a_{22} F_2 + \cdots + a_{2m} F_m \qquad + a_2U_2 \\ \cdot \quad \cdot\cdot \quad \cdot\cdot \qquad \cdot\cdot \qquad \cdot\cdot \quad \cdot \\ z_n = a_{n1}F_1 + a_{n2} F_2 + \cdots + a_{nm} F_m \qquad\qquad + a_nU_n \end{array} \tag{17}$$

The coefficients of the various factors in the equations above are estimated by the factor analysis technique and become what is known as the "factor pattern" or simply "pattern". In a pattern the common factors (F_j) may be correlated or uncorrelated, however the unique factors (U_j) are always assumed to be uncorrelated among themselves and with all common factors.

Factor analysis yields not only patterns but also correlations between the variables and the factors. A table of such correlations is called a "factor structure". Both a pattern and a structure are necessary for a complete solution.

PRINCIPAL-FACTOR SOLUTION

Derivation

Although several solutions are available to solve the factor analysis problem, only the principal-factor solution will be considered in this paper. The derivation of the principal-factor solution was taken from Hotelling (1933) and Harman (1962). The principal-factor solution was not used to any great extent until the development of high speed computers, because of the large amount of time required to complete the solution by hand methods.

The analysis is begun with a factor F_1 whose contribution to the communalities of the variables has as great a total as possible, then the first factor residual correlations are obtained. Next, a second factor F_2, independent of F_1, with a maximum contribution to the residual communality is found. This process is continued until the total communality is analyzed.

The factor pattern to be determined may be represented by

$$z_j = a_{j1} F_1 + \cdot \cdot \cdot + a_{jp} F_p + \cdot \cdot \cdot + a_{jm} F_m \; (j = 1, 2, \cdot \cdot \cdot \cdot , n) \qquad (18)$$

where the unique factor has been omitted. The sum of squares of the factor coefficients gives the communality of a particular variable, while any term a_{jp}^2 indicates the contribution of the factor F_p to the communality of z_j. The first stage of the principal-factor method involves the selection of the first factor coefficients a_{j1} so as to make the sum of the contributions of that factor to the total communality a maximum. This sum is given by

$$V_1 = a_{11}^2 + a_{21}^2 + \cdots + a_{n1}^2 , \tag{19}$$

and the coefficients a_{j1} must be chosen to make V_1 a maximum under the conditions

$$r_{jk} = \sum_{p=1}^{m} a_{jp}\, a_{kp} \qquad (j, k = 1, 2, \ldots , n) \tag{20}$$

where $r_{jk} = r_{kj}$ and r_{jj} is the communality h_j^2 of variable z_j. The conditions (20) say that the reproduced correlations are to be replaced by the observed correlations, implying the assumption of zero residuals.

In order to maximize a function of n variables when the variables are connected by an arbitrary number of auxilary equations, the method of Lagrange multipliers is particularly well adapted (Osgood, 1932). This method is employed to maximize V_1, which is a function of the n variables a_{j1} under the $\frac{1}{2}n(n + 1)$ conditions (20) among all coefficients a_{jp}. Let

$$2T = V_1 - \sum_{j,k=1}^{n} \mu_{jk}\, r_{jk} = V_1 - \sum_{j,k=1}^{n} \sum_{p=1}^{m} \mu_{jk}\, a_{jp}\, a_{kp} , \tag{21}$$

where $\mu_{jk} (= \mu_{kj})$ are the Lagrange multipliers. The partial derivative of this new function T with respect to any one of the n variables a_{j1} is set equal to zero, namely

$$\frac{\partial T}{\partial a_{j1}} = a_{j1} - \sum_{k=1}^{n} \mu_{jk} a_{k1} = 0, \tag{22}$$

and similarly put the partial derivative with respect to any of the other coefficients a_{jp} $(p \neq 1)$ equal to zero, that is,

$$\frac{\partial T}{\partial a_{jp}} = - \sum_{k=1}^{n} \mu_{jk} a_{kp} = 0 \qquad (p \neq 1) . \tag{23}$$

The two sets of equations (22) and (23) may be combined as follows:

$$\frac{\partial T}{\partial a_{jp}} = \delta_{1p} a_{j1} - \sum_{k=1}^{n} \mu_{jk} a_{kp} = 0 \; (p = 1, 2, . . . , m) \tag{24}$$

where the Kronecker $\delta_{1p} = 1$ if $p = 1$ and $\delta_{1p} = 0$ if $p \neq 1$.

Multiply (24) by a_{j1} and sum with respect to j, obtaining

$$\delta_{1p} \sum_{j=1}^{n} a_{j1}^{2} - \sum_{j=1}^{n} \sum_{k=1}^{n} \mu_{jk} a_{j1} a_{kp} = 0. \tag{25}$$

Now, the expression $\sum_{j=1}^{n} \mu_{jk} a_{j1}$ is equal to a_{k1} according to (22) and setting $\sum_{j=1}^{n} a_{j1}^{2} = \lambda_1$, equation (25) may be written as follows:

$$\delta_{1p} \lambda_1 - \sum_{k=1}^{n} a_{k1} a_{kp} = 0 \quad . \tag{26}$$

Upon multiplying (26) by a_{jp} and summing for p, this equation becomes

$$a_{j1} \lambda_1 - \sum_{k=1}^{n} a_{k1} \sum_{p=1}^{m} a_{jp} a_{kp} = 0 \tag{27}$$

or, upon applying the conditions from (20)

$$\sum_{k=1}^{n} r_{jk} a_{k1} - \lambda_1 a_{j1} = 0 . \tag{28}$$

The n equations may be rewritten in full as follows :

$$\begin{aligned}
&(h_1^2 - \lambda) a_{11} + r_{12} a_{21} + r_{13} a_{31} + . . . + r_{1n} a_{n1} = 0, \\
&r_{21} a_{11} + (h_2^2 - \lambda) a_{21} + r_{23} a_{31} + . . . + r_{2n} a_{n1} = 0, \\
&r_{31} a_{11} + r_{32} a_{21} + (h_3^2 - \lambda) a_{31} + . . . + r_{3n} a_{n1} = 0 \\
&\quad . . \qquad . . \qquad\quad . \qquad\quad . \quad . \\
&r_{n1}a_{11} + r_{n2} a_{21} + r_{n3} a_{31} + . . . + (h_n^2 - \lambda) a_{n1} = 0
\end{aligned} \tag{29}$$

where the parameter of (28) is designated by λ without a subscript.

Thus, the maximization of (19) under the conditions (20) leads to a system of n equations (29) for the solution of the n unknowns a_{j1}. A necessary and sufficient condition for this system of n homogeneous equations to have a non-trivial solution is the vanishing of the determinant of coefficients of the a_{j1}. This condition may be written

$$\begin{vmatrix} (h_1^{\,2} - \lambda) & r_{12} & r_{13} & \cdots & r_{1n} \\ r_{21} & (h_2^{\,2} - \lambda) & r_{23} & \cdots & r_{2n} \\ r_{31} & r_{32} & (h_3^{\,2} - \lambda) & \cdots & r_{3n} \\ \cdot\,\cdot & \cdot\,\cdot & \cdot\,\cdot & & \cdot\,\cdot \\ r_{n1} & r_{n2} & r_{n3} & \cdots & (h_n^{\,2} - \lambda) \end{vmatrix} = 0 \quad (30)$$

If the determinant in (30) were expanded it would lead to an n-order polynomial in λ. This equation is known as a characteristic equation. Some of the important properties of characteristic equations which apply to factor analysis include the fact that all roots are real and that a q-fold multiple root substituted for λ in (30) reduces the rank of the determinant to $(n - q)$.

When a simple root of the characteristic equation is substituted for λ in (29) a set of homogeneuous linear equations of rank $(n - 1)$ is obtained. This set of equations has a family of solutions, all of which are proportional to one particular solution. It follows that the factor of proportionality is $\lambda_1 = \sum_{j=1}^{n} a_{j1}^{\,2}$. This expression is precisely V_1, the quantity which is to be maximized. In other words, V_1 is equal to the largest root λ_1 of the characteristic equation (30).

The problem of finding the coefficients a_{j1} of the first factor F_1, which will account for as much of the total communality as possible, is then solved. The largest root λ_1 of (30) is substituted in (29), and any solution $\alpha_{11}, \alpha_{21}, \ldots, \alpha_{n1}$ is obtained. These values are divided by the square root of the sum of their squares and then multiplied by $\sqrt{\lambda_1}$ to satisfy relation (19). The resulting quantities are

$$a_{j1} = \alpha_{j1} \sqrt{\lambda_1} \Big/ \sqrt{(\alpha_{11}^2 + \alpha_{21}^2 + \ldots + \alpha_{n1}^2)} \quad (j = 1, 2, \ldots, n) \quad (31)$$

which are the desired coefficients of F_1 in the factor pattern (18).

The roots (λ's) of a characteristic equation (30) are referred to as "eigenvalues". The solution to the set of equations (29) corresponding to each eigenvalue leads to a vector which is called an "eigenvector". The mathematical problem can be expressed in the form: Find a number λ and an n dimensional vector $\underline{x} \neq 0$ such that

$$R\,\underline{x} = \lambda\,\underline{x} \quad (32)$$

where R is the correlational matrix. Any number λ_p satisfying this equation is called an eigenvalue of R and its associated vector $\underline{x}_p = \left[\alpha_{1p}, \alpha_{2p}, \ldots \alpha_{np}\right]$ is called an eigenvector of R. An eigenvector scaled to (31) is designated $\underline{a}_p = \left[a_{1p}, a_{2p}, \ldots, a_{np}\right]$.

Having determined the coefficients a_{j1}, of the first factor F_1, the next problem is to find a factor which will account for a maximum of the residual communality. In order to do this, it is necessary to obtain the first-factor residual correlations. Furthermore, in obtaining still other factors the residual correlations with two, three, . . . , $(m - 1)$ factors removed are employed, and hence a suitable notation is required. A convenient notation for the residual correlation of r_{jk} with s factors removed is ${}_{s}r_{jk}$. Thus, when the first factor has been obtained the first factor residuals become

$$ {}_1r_{jk} = r_{jk} - a_{j1}\, a_{k1} = a_{j2}\, a_{k2} + a_{j3}\, a_{k3} + \ldots + a_{jm}\, a_{km} \; . \qquad (33)$$

More generally, the matrix of first-factor residuals may be expressed in matrix form by:

$$R_1 = R - Q_1 \qquad (34)$$

where

$$Q_1 = \underline{a}_1\, \underline{a}_1{}' \qquad (35)$$

represents the n x n symmetric matrix of products of the first-factor coefficients, which appear in the column vector $\underline{a}_1$.

In determining the coefficients of the second factor F_2, it is necessary to maximize the quantity

$$V_2 = a_{12}{}^2 + a_{22}{}^2 + \ldots + a_{n2}{}^2 \qquad (36)$$

which is the sum of the contributions of F_2 to the residual communality. This maximization is subject to the conditions (33) which are analogous to the restrictions (20) in the case of the first factor. The theory of characteristic equations provides the basis for determining the coefficients of the second and subsequent factors. It is not necessary, however, to carry through an analysis for maximizing the contributions of F_2 to the residual communality. Instead it will be shown that the required maximum eigenvalue of R_1 is, in fact, the second largest eigenvalue of the original correlation matrix R.

If a_p represents the m eigenvectors of R (properly scaled), it can be determined whether they are also eigenvectors of R_1. Postmultiplying the matrix R_1 by any vector a_p yields

$$R_1 \, a_p = (R - \underline{a}_1 \, \underline{a}_1') \, a_p \tag{37}$$

from the definition (34) of the residual matrix. Expanding this expression and applying (32) produces:

$$R_1 \, \underline{a}_p = R\underline{a}_p - \underline{a}_1 \, \underline{a}_1' \, \underline{a}_p = \lambda_p \, \underline{a}_p - \underline{a}_1 \, \underline{a}_1' \, \underline{a}_p \, . \tag{38}$$

Now consider the two cases: $p = 1$ and $p \neq 1$. a) When $p = 1$, $\lambda_1 = \underline{a}_1' \, \underline{a}_1$ according to (38), so that the above expression reduces to

$$R_1 \, \underline{a}_1 = 0. \tag{39}$$

In other words, the eigenvector corresponding to the largest eigenvalue λ_1 of R is also an eigenvector of R_1 but its associated eigenvalue in R_1 is zero. b) When $p \neq 1$, $\underline{a}_1' \, \underline{a}_p = 0$ according to (26) and expression (38) becomes:

$$R_1 \, \underline{a}_p = \lambda_p \, \underline{a}_p - \underline{a}_1 \cdot 0 = \lambda_p \, \underline{a}_p \tag{40}$$

which says that except for λ_1, the eigenvalues of R_1 are identical with those of R and their associated eigenvectors are also identical. The expressions (34) and (40) prove that the eigenvectors of R_1 are identical with those of R, and that they have corresponding eigenvalues except that corresponding to the eigenvector $\underline{a}_1$ in R_1 is a zero eigenvalue in place of λ_1 in R.

From the foregoing it is clear that λ_2 of R is the largest eigenvalue of R_1. In other words, to obtain the coefficients of the second factor F_2 from the largest eigenvalue of the residual matrix R_1 it suffices to extract the second largest eigenvalue of the original matrix R.

By the same type of argument, the successive eigenvalues and their associated eigenvectors are obtained directly from the original correlation matrix R, until m factors have been extracted.

When unities are placed in the diagonal of R then usually m = n. If some numbers less than unities (estimates of communalities) are placed in the diagonal, and the positive semi-definite property of R is preserved, then m will be less than n, and all eigenvalues will be real and nonnegative. When negative eigenvalues occur in the course of the computation, it is evidence that the requirement of positive semi-definiteness has been violated.

An important mathematical property of the principal-factor solution is that of orthogonality of the column vectors of factor coefficients. This property is expressed by

$$\underline{a}_p' \, \underline{a}_q = \delta_{pq} \, \lambda_p \tag{41}$$

or in expanded algebraic form:

$$\sum_{j=1}^{n} a_{jp}^{2} = \lambda_p$$

$$(p, q=1, 2, . . . , m; p \neq q).(42)$$

$$\sum_{j=1}^{n} a_{jp} \, a_{jq} = 0.$$

Computational Procedures

The method used in deriving the principal-factor solution does not lend itself to efficient computation of the factor problem. The method that will be presented in this paper is taken from Hotelling (1936) and Harman (1962).

This method involves an iterative scheme which yields a root of the characteristic equation and the coefficients of the associated factor simultaneously. The roots appear in descending order of magnitude upon successive applications of the method. For this reason the method is especially suitable in practical situations where only a few of the largest characteristic roots and associated factor coefficients are required. From a practical point of view it is desirable if a small number of roots will account for the total communality and this is one of the chief aims of factor analysis.

The iterative process is begun by selecting an arbitrary set of n numbers, and transforming them again and again by use of the observed correlations until they converge to the desired coefficients of the first principal factor. Thus, an arbitrary set $[\alpha_{11}, \alpha_{21}, \ldots, \alpha_{n1}]$ is transformed into a new set $[\gamma_{11}, \gamma_{21}, \ldots, \gamma_{n1}]$ as follows:

$$\gamma_{j1} = \sum_{k=1}^{n} r_{jk}\,\alpha_{k1} \qquad (j = 1, 2, \ldots, n) \tag{43}$$

or in matrix notation:

$$\underline{\gamma} = R\,\underline{\alpha} \tag{44}$$

where $\underline{\alpha}$ and $\underline{\gamma}$ are the original and transformed n x 1 column vectors, respectively, and R is the n x n correlation matrix.

If the number α_{j1} are proportional to the directional cosines of any line through the origin, then the numbers γ_{j1} are proportional to the direction cosines of a new line resulting from the rotation (43). In general, the line associated with the numbers γ_{j1} is distinct from the line corresponding to the α_{j1}.

In practice it cannot be expected that the arbitrary numbers α_{j1} will be so selected as to be proportional to the direction cosines of one of the principal axes. The iterative process then involves the use of the derived numbers γ_{j1} as a new set of arbitrary numbers in place of α_{j1} and the transformation (44) becomes

$$\underline{\xi} = R\underline{\gamma} \tag{45}$$

So again, if the numbers γ_{j1} are proportional to the direction cosines of any line through the origin, then the numbers ξ_{j1} will be proportional to the direction cosines of a new line corresponding to the original one under transformation (45). This process is continued until the ratios among the quantities obtained at any stage converge to the corresponding ratios among the coefficients of F_1 to any specified degree of accuracy. The proof of the convergence of these ratios to those of the coefficients a_{j1} of the first principal factor is given by Hotelling (1933). A convenient procedure is to divide each of the trial values by a fixed one of them, say the largest. Then the next value obtained, corresponding to this number, will be an approximation to the characteristic root λ_1.

Instead of calculating the successive values γ_{j1}, ξ_{j1} etc., and substituting them in equations like (44) and (45), a modification will next be introduced which greatly accelerates convergence (Hotelling, 1936). This simplification is accomplished by the formal algebraic substitution of the successive estimates in these equations. Upon substituting the values for γ_{j1} from (44) into the right hand member of (45), the latter equation takes the form

$$\underline{\xi} = R(R\underline{\alpha}) = R^2\underline{\alpha} \, . \tag{46}$$

Consequently, if the correlation matrix is first squared and then used for the transformation of the original set of arbitrary numbers α_{j1}, the same results are obtained as that accomplished by the transformation (44) followed by the transformation (45) and effectively does the job of two iterations. The improvement in the iteration process need not end with the employment of R^2 as higher powers of R will increase the speed of convergence.

The second and remaining principal factors may be determined by the same method, and convergence can be accelerated by the use of a convenient power of the matrix of residual correlations. It is not necessary, however, to obtain this power of the residual matrix by repeated squarings, as was done in the case of the original matrix of correlations. Instead, the determination already made of the power of R and the following algebraic properties of matrices can be employed for this purpose.

In getting the square of the residual matrix algebraically from (34)

$$R_1{}^2 = R^2 - 2\,R\,Q_1 + Q_1{}^2 \qquad (47)$$

the terms RQ_1 and $Q_1{}^2$ appear. If these terms can be expressed by quantities already known then the actual squaring of R_1 will be eliminated. Thus, from the definition (35) of the product matrix,

$$\begin{aligned} Q_1{}^2 &= (\underline{a}_1\underline{a}_1{}')\ (\underline{a}_1\underline{a}_1{}') \\ &= \underline{a}_1(\underline{a}_1{}'\underline{a}_1)\ \underline{a}_1{}' \\ &= \underline{a}_1\ \lambda_1\ \underline{a}_1{}' \qquad\qquad \text{From (41)} \\ &= \lambda_1\ \underline{a}_1\ \underline{a}_1{}' \qquad \text{since } \lambda_1 \text{ is a scalar and} \end{aligned}$$

applying definition (35) to the last expression produces:

$$Q_1^{\,2} = \lambda_1 Q_1 \quad . \tag{48}$$

Also, from the matrix formulation of (28) for the particular case of the first principal factor of the correlation matrix R, the following relationship

$$R\underline{a}_1 = \lambda_1 \underline{a}_1 \tag{49}$$

provides the basis for expressing RQ_1 in terms of known quantities. Postmultiplying (49) by $\underline{a}_1{}'$ and applying definition (35) yields:

$$RQ_1 = \lambda_1 Q_1 \tag{50}$$

Upon substituting the known quantities from (48) and (50) for the terms RQ_1 and $Q_1^{\,2}$ in (47), the square of the residual matrix becomes:

$$\begin{aligned} R_1^{\,2} &= R^2 - 2\lambda_1 Q_1 + \lambda_1 Q_1 \\ &= R^2 - \lambda_1 Q_1 \quad . \end{aligned} \tag{51}$$

In other words, actual squaring of the residual matrix is not necessary since the square of the correlation matrix and the matrix of products of first factor coefficients are available.

The preceding analysis can be generalized to higher powers of the residual matrix. For example upon squaring (51)

$$R_1^{\,4} = R^4 - 2\,\lambda_1\,R^2\,Q_1 + \lambda_1^{\,2}\,Q_1^{\,2} \tag{52}$$

and successively applying (50) to the middle term and (48) to the last term on the right, there results:

$$R_1{}^4 = R^4 - 2\,\lambda_1{}^2\,R\,Q_1 + \lambda_1{}^3\,Q_1$$

$$= R^4 - 2\,\lambda_1{}^3\,Q_1 \quad + \lambda_1^3\,Q_1$$

$$= R^4 - \lambda_1{}^3\,Q_1 \quad . \tag{53}$$

In similar fashion, it can be shown that for any positive integer e,

$$R_1{}^e = R^e - \lambda_1{}^{e-1}\;Q_1 \quad . \tag{54}$$

Therefore the e power of the residual matrix is expressed in terms of the e power of the original correlation matrix, eliminating the actual multiplications of the residual matrix.

A NUMERICAL EXAMPLE

First-Factor Coefficients

The data used in this example was taken from a report on the evaluation of flight performance in the United States Air Force B-47 training program (Woolman, 1955). The student pilots were required to perform six specific tasks on the B-47 simulator trainer. A numerical score was recorded by the instructor based on the students performance of the task. In order to study the relationships between the six tasks, a sample of 76 students was selected and intercorrelations between all tasks were computed. These correlations are presented in Table 1. This is the basic data that will be used to illustrate the principal - factor solution of the factor analysis problem.

Table 1. Intercorrelation matrix

Task	1	2	3	4	5	6
1. Turns	-	.78	.01	.17	.09	.29
2. Change of A/S		-	.18	.33	.19	.26
3. ADF			-	.54	.63	.31
4. VOR				-	.40	.23
5. ILAS					-	.36
6. GCA						-

The first step in solving the factor analysis problem is to decide how to estimate the communalities to be entered on the diagonal of the matrix. Several suggestions have been offered on how to estimate these numbers or communalities (Harmon, 1962). One of the simplest methods is to choose the highest correlation in each row or column and use this as the estimate of the jth communality. Using this procedure the reduced correlation matrix is given in Table 2 with the elements above and below the diagonal included for computational purposes.

Table 2. Reduced correlation matrix: R

	1	2	3	4	5	6	S_j	$\alpha_{j1}^{(1)}$
1	.78	.78	.01	.17	.09	.29	2.12	.8413
2	.78	.78	.18	.33	.19	.26	2.52	1.0000
3	.01	.18	.63	.54	.63	.31	2.30	.9127
4	.17	.33	.54	.54	.40	.23	2.21	.8770
5	.09	.19	.63	.40	.63	.36	2.30	.9127
6	.29	.26	.31	.23	.36	.36	1.81	.7182

The sums of the rows of R are entered in the column S_j. The set of trial values $\alpha_{j1}^{(1)}$ are obtained by dividing each value of S_j by the largest value appearing in the column. In the above case, 2.52 was the divisor used in obtaining the $\alpha_{j1}^{(1)}$ set of trial values. The higher powers of the matrix R were then computed to increase the speed of convergence.

Table 3. Square of correlation matrix: R^2

	1	2	3	4	5	6	$S_j^{(2)}$	$T_j^{(2)}$	$\alpha_{j1}^{(2)}$
1	1.34	1.37	.39	.59	.45	.61	4.75	4.75	.8377
2	1.37	1.46	.64	.80	.68	.72	5.67	5.67	1.0000
3	.39	.64	1.21	1.02	1.16	.71	5.13	5.13	.9048
4	.59	.80	1.02	.93	.97	.65	4.96	4.96	.8748
5	.45	.68	1.16	.97	1.13	.72	5.11	5.10	.8995
6	.61	.72	.71	.65	.72	.56	3.97	3.97	.7002

The column headed T_j has been added to Table 3 for checking purposes. This check is made by computing the product of R by the column of values S_j from Table 2. These values

$$T_j^{(2)} = \sum_{k=1}^{6} r_{kj} S_k \qquad (j = 1, 2, . . . , 6) \tag{55}$$

should agree, except for rounding errors, with the respective sums S_j of the rows of R. It is advantageous to compute the column T_j first since this can be done without actually squaring the matrix and the next set of values α_{j1} can be computed. In this manner it is possible to make one further comparison to ascertain whether there is sufficient agreement between the last two sets of trial values α_{j1} to discontinue raising the matrix R to any higher power. The second set of trial values $\alpha_{j1}^{(2)}$ are computed in exactly the same manner as the values $\alpha_{j1}^{(1)}$. Since the agreement between $\alpha_{j1}^{(1)}$ and $\alpha_{j1}^{(2)}$ was not too good, the matrix R was raised to a higher power.

Table 4. Fourth power of correlation matrix: R^4

	1	2	3	4	5	6	$S_j^{(4)}$	$T_j^{(4)}$	$\alpha_{j1}^{(4)}$
1	4.75	5.30	3.43	3.67	3.51	3.13	23.79	23.78	.8385
2	5.30	6.04	4.36	4.50	4.41	3.75	28.36	28.36	1.0000
3	3.43	4.36	4.92	4.51	4.83	3.45	25.50	25.48	.8984
4	3.67	4.50	4.51	4.26	4.46	3.33	24.73	24.71	.8713
5	3.51	4.41	4.83	4.46	4.75	3.44	25.40	25.38	.8949
6	3.13	3.75	3.45	3.33	3.44	2.65	19.75	19.74	.6960

Table 5. Eighth power of correlation matrix: R^8

	1	2	3	4	5	6	$S_j^{(8)}$	$T_j^{(8)}$	$\alpha_{j1}^{(8)}$
1	98.00	115.87	100.58	98.46	100.42	79.17	592.5	592.5	.8393
2	115.87	137.34	120.50	117.62	120.22	94.37	705.9	705.9	1.0000
3	100.58	120.50	110.55	106.64	109.96	84.84	633.1	633.1	.8968
4	98.46	117.62	106.64	103.19	106.15	82.27	614.3	614.3	.8702
5	100.42	120.22	109.96	106.15	109.38	84.50	630.6	630.6	.8930
6	79.17	94.37	84.84	82.27	84.50	65.71	490.9	490.9	.6953

Table 6. Sixteenth power of correlation matrix: R^{16}

	1	2	3	4	5	6	$S_j^{(16)}$	$T_j^{(16)}$	$\alpha_{j1}^{(16)}$
1	-	-	-	-	-	-	----	366210	.8394
2	-	-	-	-	-	-	----	436283	1.0000
3	-	-	-	-	-	-	----	391144	.8965
4	-	-	-	-	-	-	----	379596	.8701
5	-	-	-	-	-	-	----	389644	.8931
6	-	-	-	-	-	-	----	303319	.6952

Note by using equation (55) it was not necessary to actually square the eighth power of matrix R in order to get the set of trial values $\alpha_{j1}^{(16)}$ in Table 6. The differences between the successive sets of trial values are summarized in Table 7.

Table 7. Differences between successive trial values

Variable	$\alpha_{j1}^{(2)} - \alpha_{j1}^{(1)}$	$\alpha_{j1}^{(4)} - \alpha_{j1}^{(2)}$	$\alpha_{j1}^{(8)} - \alpha_{j1}^{(4)}$	$\alpha_{j1}^{(16)} - \alpha_{j1}^{(8)}$
1	-.0036	.0008	.0008	.0001
2	0	0	0	0
3	-.0079	-.0064	-.0016	-.0003
4	-.0022	-.0035	-.0011	-.0001
5	-.0132	-.0046	-.0019	.0001
6	-.0180	-.0042	-.0007	-.0001

The $\alpha_{j1}^{(8)}$ values agree with the $\alpha_{j1}^{(16)}$ values to three decimal places. The next step is to calculate the F_1 coefficients a_{j1}. The calculation of these values using equation (31) is given in Table 8.

Table 8. Calculation of the F_1 coefficients

Variable	$\alpha_{j1}^{(16)}$	First Iteration		/ /	Third Iteration		Fourth Iteration		a_{j1}
		γ_{j1}	α_{j1}	/ /	γ_{j1}	α_{j1}	γ_{j1}	α_{j1}	
1	.8394	1.8736	.8388		1.8730	.8389	1.8730	.8389	.588
2	1.0000	2.2337	1.0000		2.2328	1.0000	2.2327	1.0000	.701
3	.8965	2.0012	.8959		1.9995	.8955	1.9993	.8955	.628
4	.8701	1.9438	.8702		1.9425	.8700	1.9424	.8700	.610
5	.8931	1.9913	.8915		1.9900	.8913	1.9894	.8910	.625
6	.6952	1.5533	.6954		1.5522	.6952	1.5521	.6952	.488

$$\lambda_1 = 2.2327 \qquad \sum_{j=1}^{6} \alpha_{j1}^2 = 4.5398 \qquad \sqrt{\lambda_1} \Big/ \sqrt{\sum \alpha_{j1}^2} = .7013$$

The γ_{j1} values in each iteration were computed using equation (43). After the γ_{j1} values had been computed the α_{j1} values for that iteration were computed by dividing each γ_{j1} by the largest value of γ_{j1}. After four iterations the successive sets of α_{j1} values agreed to three decimal places and were considered stable enough to calculate the a_{j1} coefficients using equation (31).

Second - Factor Coefficients

It was shown that the matrix of first-factor residuals R_1 is given by the equation

$$R_1 = R - Q_1 \tag{34}$$

where

$$Q_1 = \underline{a}_1 \, \underline{a}_1' \quad . \tag{35}$$

In this example the matrix Q_1 is obtained by multiplying the column vector $\underline{a}_1$ by its transpose $\underline{a}_1'$ where

$$\underline{a}_1 = \begin{bmatrix} ,588 \\ .701 \\ .628 \\ .610 \\ .625 \\ .488 \end{bmatrix} \quad .$$

Table 9. Product matrix: $Q_1 = \underline{a}_1 \underline{a}_1'$

	1	2	3	4	5	6	E_{j1}	a_jD_1
1	.346	.412	.369	.359	.368	.287	2.141	2.14
2	.412	.491	.440	.428	.438	.342	2.551	2.55
3	.369	.440	.394	.383	.392	.306	2.284	2.29
4	.359	.428	.383	.372	.381	.298	2.221	2.22
5	.368	.438	.392	.381	.391	.305	2.275	2.27
6	.287	.342	.306	.298	.305	.238	1.776	1.77

In order to check the calculation of the elements of the product matrix Q_1, obtain the sums of the rows E_{j1} and compare them with the corresponding values a_jD_1 where

$$D_1 = \sum_{k=1}^{6} a_{k1} \quad . \tag{56}$$

The sum of the first-factor coefficients D_1 for this example is 3.640.

The matrix R_1 is computed by subtracting the elements in matrix Q_1 from the corresponding elements in matrix R.

Table 10. Matrix of first-factor residuals: R_1

	1	2	3	4	5	6	S_{j1}	$a_{j2}^{(1)}$
1	.434	.368	-.359	-.189	-.278	.003	-.021	-.6176
2	.368	.289	-.260	-.098	-.248	-.082	-.031	-.9118
3	-.359	-.260	.236	.157	.238	.004	.016	.4706
4	-.189	-.098	.157	.168	.019	-.068	-.011	-.3235
5	-.278	-.248	.238	.019	.239	.055	.025	.7353
6	.003	-.082	.004	-.068	.055	.122	.034	1.0000

Since the S_{j1} sums were very small the $\alpha_{j2}^{(1)}$ set of trial values were used to compute the second factor coefficients rather than raising the matrix R_1 to a higher power. The results of these calculations are given in Table 11.

Table 11. Calculation of the F_2 coefficients

Variable	$\alpha_{j2}^{(1)}$	First Iteration		/ /	Fourth Iteration		Fifth Iteration		
		γ_{j2}	α_{j2}	/ /	γ_{j2}	α_{j2}	γ_{j2}	α_{j2}	a_{j2}
1	-.6176	-.9128	-1.0000		-1.2559	-1.0000	-1.2549	-1.0000	-.667
2	-.9118	-.8458	-0.9266		-1.0036	-0.7991	-1.0028	-0.7991	-.535
3	.4706	.6981	0.7647		0.9626	0.7665	0.9618	0.7664	.513
4	-.3235	.1716	0.1880		0.4570	0.3639	0.4567	0.3639	.244
5	.7353	.7344	0.8046		0.8273	0.6587	0.8267	0.6588	.441
6	1.0000	.2592	0.2840		0.0855	0.0681	0.0854	0.0680	.046
$\lambda_2 =$	1.2549	$\sum_{j=1}^{6} \alpha_{j2}^2 = 2.7970$				$\sqrt{\lambda_2} \Big/ \sqrt{\sum \alpha_{j2}^2} = .6698$			

Five iterations were required before the α_{j2} set of trial values were considered stable enough to compute the a_{j2} coefficients. In order to determine whether two factors accounted for most of the communality, it was necessary to compute the matrix of the second-factor residuals R_2 where

$$R_2 = R_1 - Q_2 \tag{57}$$

and

$$Q_2 = \underline{a}_2 \, \underline{a}_2' \quad . \tag{58}$$

In this case

$$\underline{a}_2 = \begin{bmatrix} -.667 \\ -.535 \\ .513 \\ .244 \\ .441 \\ .046 \end{bmatrix} .$$

Table 12. Product matrix: $Q_2 = \underline{a}_2 \, \underline{a}_2'$

	1	2	3	4	5	6	E_{j2}	$a_j D_2$
1	.445	.356	-.342	-.163	-.294	-.031	-.029	-.028
2	.356	.286	-.274	-.131	-.236	-.025	-.024	-.022
3	-.342	-.274	.263	.125	.226	.024	.022	.022
4	-.163	-.131	.125	.060	.108	.011	.010	.010
5	-.294	-.236	.226	.108	.194	.020	.018	.019
6	-.031	-.025	.024	.011	.020	.002	.001	.002

The check procedure for the matrix Q_2 is analagous to that of the product matrix Q_1.

Table 13. Second factor residuals: R_2

	1	2	3	4	5	6
1	-.011					
2	.012	.003				
3	-.017	.014	-.027			
4	-.026	.033	.032	.108		
5	.016	-.012	.012	-.089	.045	
6	.028	-.057	-.020	-.077	.035	.120

Two factors have accounted for the greater part of the communality, since the largest value in the second-factor residual matrix R_2 is .120. In addition the majority of the entries have an absolute value of less than .05. For practical purposes and for the purposes of this paper, the two factors are considered an adequate solution to this factor problem. The results are summarized in Table 14.

Table 14. Principal-factor pattern for six B-47 simulator tasks

Variable		Pattern Coefficients		Communality	
j		F_1	F_2	Original	Calculated
1.	Turns	.588	−.667	.78	.791
2.	Change of A/S	.701	−.535	.78	.778
3.	ADF	.628	.513	.63	.658
4.	VOR	.610	.244	.54	.432
5.	ILAS	.625	.441	.63	.585
6.	GCA	.488	.046	.36	.240
Total				3.72	3.484
Contribution of factor (V_p)		2.232	1.250		
% of total original communality		60.0	33.6		93.6

Although 93.6% of the estimated communality has been accounted for after the extraction of two factors, we do not know whether the original estimates of the communalities were correct. One method of refining the estimates of communalities is an iteration technique (Harman, 1962). This method involves

obtaining a principal-factor solution such as that presented in Table 14, then refactoring with the calculated communalities as the next estimates. This process is continued until agreement between successive sets of communalities has been achieved to any desired degree of accuracy. This technique is practical if a high speed digital computer is available. The solution presented in Table 14 indicates some deviations between the estimates of the communalities and the calculated communalities. In actual practice these deviations are large enough to warrant getting better estimates of the communalities.

Interpretation

Assuming that the solution above is adequate, the next step is to interpret the results of this analysis. The factor pattern as it appears in Table 14 indicates that there is 1) a general factor where all six variables are contributing about equally and 2) a second factor in which variables 1 and 2 are contributing something quite different from variables 3, 4 and 5. This is evident by the negative loadings on variables 1 and 2 and positive loadings on variables 3, 4 and 5. The pattern definitely indicates the presence of two factors.

It is often possible to simplify the interpretation of the original factor pattern by making an orthogonal rotation in the factor space. The purpose of the orthogonal rotation is to obtain "simple structure", (Thurston, 1947). A simple structure may be viewed as an attempt to reduce the complexity of the original factors.

The original factor pattern has been plotted in Figure 1. The axes were rotated 50 degrees in a clockwise direction in an attempt to achieve

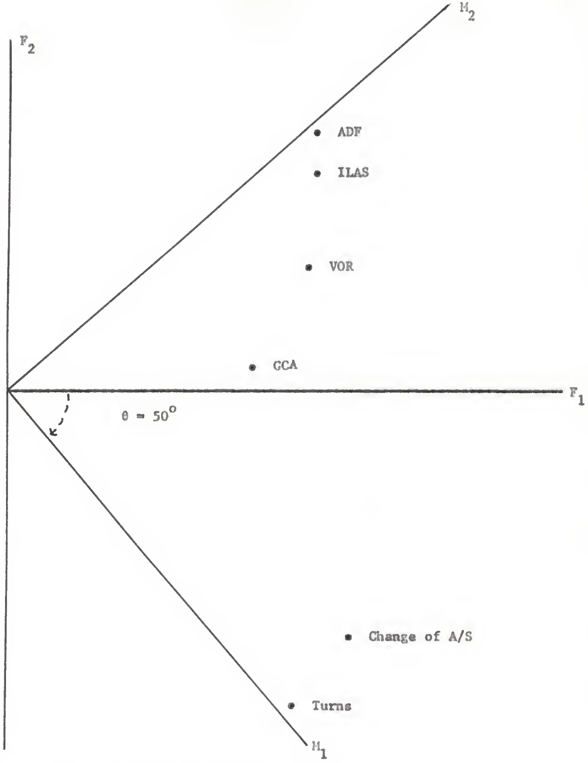

Figure 1. Orthogonal rotation of axes.

high loadings on tasks 3 - 6 in one factor and high loadings on tasks 1 and 2 in the second factor. The required transformation is accomplished by the following set of equations:

$$b_1 = \underline{a}_1 \text{ Cos } \theta + \underline{a}_2 \text{ Sin } \theta$$

$$b_2 = -\underline{a}_1 \text{ Sin } \theta + \underline{a}_2 \text{ Cos } \theta$$

where θ is the angle of rotation. This transformation can also be expressed in matrix form

$$M = C \begin{bmatrix} \text{Cos } \theta & -\text{Sin } \theta \\ \text{Sin } \theta & \text{Cos } \theta \end{bmatrix}$$

where C is the pattern matrix from Table 14 and M is the rotated matrix pattern.

In this example

Sin (-50) = -.766 Cos (-50) = .643

and

$$M = \begin{bmatrix} .588 & -.667 \\ .701 & -.535 \\ .628 & .513 \\ .610 & .244 \\ .625 & .441 \\ .488 & .046 \end{bmatrix} \begin{bmatrix} .643 & .766 \\ -.766 & .643 \end{bmatrix} = \begin{bmatrix} .889 & .022 \\ .861 & .193 \\ .011 & .811 \\ .205 & .624 \\ .064 & .762 \\ .278 & .403 \end{bmatrix} .$$

It is evident that some simplification was achieved by this rotation. Variables 1 and 2 are isolated in one factor while variable 3 - 6 are highly loaded in the other factor. The interpretation of these results seems to depend upon the relative complexity of the six variables or tasks. The two tasks, Turns and Change of Airspeed, were simple manuevers and were highly correlated consequently they make up one of the two factors. The remaining four tasks were quite complex in contrast and tended to group together in a second factor.

The same data was factor analyzed in 1955 by the writer of this report. The centroid solution was used to determine the factor pattern in the earlier analysis (Guilford, 1954). An orthoganal rotation of the axes was made to give a clearer interpretation of the results. The rotated factors of the two solutions are presented in Table 15 for comparison.

Table 15. Comparison of two factor solutions

Variable		Principal - Factor		Centroid	
		M_1	M_2	M_1	M_2
1	Turns	.889	.022	.85	.04
2	Change of A/S	.861	.193	.90	.16
3	ADF	.011	.811	.00	.86
4	VOR	.205	.624	.22	.60
5	ILAS	.064	.762	.07	.76
6	GCA	.278	.403	.27	.41

The results of the two analyses agree quite closely. The slight discrepancies may be due to the failure to refine the estimates of the communalities in the principal - factor solution. It is both interesting and gratifying to this writer that the results should agree so closely using two entirely different solutions.

ACKNOWLEDGMENT

The writer wishes to express his thanks to Dr. Gary F. Krause for his helpful suggestions in the preparation of this report.

REFERENCES

Fuller, Leonard E.
Basic Matrix Theory. Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1962.

Garnett, J. C. M.
On Certain Independent Factors in Mental Measurement. Proc. Roy. Soc. Lon. 46: 91-111. 1919.

Guilford, J. P.
Psychometric Methods. New York: McGraw Hill, 1954.

Harman, Harry H.
Modern Factor Analysis. Chicago: University of Chicago Press, 1962.

Holzinger, Karl J.
Preliminary Report on Spearman-Holzinger Unitary Trait Study, Nos. 1-9. Chicago: Statistical Laboratory, Dept. of Educ., University of Chicago, 1934, 1935, 1936.

Hotelling, Harold.
Analysis of a Complex of Statistical Variables into Principal Components. Journ. Educ. Psych. 24: 417-41, 498-520. 1933.

__________.
Simplified Calculation of Principal Components. Psychometrika, 1: 27-35. 1936.

Lawley, D. N. and Maxwell, A. E.
Factor Analysis as a Statistical Method. London: Butterworths, 1963.

Osgood, William F.
Advanced Calculus. New York: Macmillan Co., 1932.

Pearson, Karl.
On Lines and Planes of Closest Fit to Systems of Points in Space. Phil. Mag. 6: 229-311. 1901.

Spearman, Charles.
General Intelligence, Objectively Determined and Measured. Amer. Journ. of Psych. 15: 201-93. 1904.

Thurstone, L. L.
Multiple Factor Analysis. Chicago: University of Chicago Press, 1947.

__________.
The Vectors of Mind. Chicago: University of Chicago Press, 1935.

Woolman, Myron.
Evaluating Flight Performance. McConnell AFB, Kansas: Training, Analysis and Development Report. 55-1, 1955.

THE USE OF FACTOR ANALYSIS IN RESEARCH

by

JAMES IVEY CLOGSTON

B. S., Friends University, 1957

AN ABSTRACT OF A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Statistics

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1964

ABSTRACT

Factor analysis is that branch of statistics which deals with the internal structure of matrices of correlations. Initially, it was developed by psychologists, with Spearman, Thomson and Thurstone pioneering the early work in the development of factor analysis. Today factor analysis is being used in the areas of geology, medicine, physical science, and many other fields besides that of psychology. This wider use of factor analysis has been brought about, primarily, by the development of high speed digital computers. The modern computer has made it possible to solve the factor problem in minutes and hours rather than the days and weeks formerly required using the desk calculator.

The chief aim of factor analysis is to describe a number of variables, say n, in terms of m factors where $m \leq n$. The model for any standardized variable say z_j can be expressed as follows

$$z_j = a_{j1} F_1 + a_{j2} F_2 + . . . + a_{jm} F_m + a_j U_j \quad (j = 1, ... , n)$$

where $F_1, F_2, ... , F_m$ represents the m factors and U_j represents the uniqueness of the jth variable. The complete mathematical model is a matrix of n such equations. The problem is to compute the coefficients of $F_1 \; F_2 \; . . . \; F_m$ for each variable j. These coefficients are called the factor pattern and represent the solution of the factor problem.

There are several solutions presented in the literature for determining a factor pattern. The principal - factor solution was chosen because of its rigorous mathematical basis. From an algebraic point of view the principal

factor solution consists of choosing a set of factors in decreasing order of their contribution to the total communality. The analysis is begun with the original correlation matrix with the communalities or estimates of the communalities as the diagonal elements. The factor F_1 whose contribution to the communalities of the variables has as great a total as possible is extracted. Then the first-factor residual correlations are obtained. A second factor F_2, independent of F_1, with maximum contribution to the residual communality is next found. This process is continued until the total communality is analyzed. The $\underline{a}_{j1}, \underline{a}_{j2}, \cdots \underline{a}_{jm}$ column vectors are the coefficients of $F_1, F_2, \cdots, F_m$ and are the desired factor pattern.

There are at least two major advantages of the principal factor solution over other available solutions:

1) No assumptions are required about the distribution of the original n variables.

2) The factors are extracted in decreasing order of size. Therefore the process can be terminated at any point and the statistician can be assured that any of the remaining factors will be smaller than the last factor extracted.

The principal-factor method is illustrated with a numerical example. A comparison of the centroid solution and the principal-factor solution of this example is presented.