CHAPTER 7
INTRODUCTION TO EXPLORATORY FACTOR ANALYSIS

From
Exploratory Factor Analysis
Ledyard R Tucker
and
Robert C. MacCallum

CHAPTER 7
INTRODUCTION TO EXPLORATORY FACTOR ANALYSIS

Factor analytic methodologies may be conceived on a continuum ranging from confirmatory techniques to pure exploratory procedures. Charles Spearman (1904 onward) was interested in confirming the idea of a general intelligence. With extended experimental evidence developed through years of studies involving larger test batteries given to larger samples of individuals, Spearman's theory of a single intellectual factor proved to be inadequate. Allowance had to be made for group factors. In the early 1930's, Thurstone broke with a common presumption based on prior assumptions as to the nature of factors and developed a general theory of multiple factor analysis. This is the theory presented in previous chapters. Thurstone's book "Vectors of Mind" (1935) presented the mathematical and logical basis for this theory. The only assumption made (and this is a question to be answered for each body of data) is that the interrelations among the measures of observed attributes may be represented by a smaller number of common factors. Needs for formal methods for checking the similarity of results from several studies have lead to development of matching procedures termed procrustes rotations. These procedures tend toward being confirmatory analyses. A return to a more general form of confirmatory factor analysis has occured with Jöreskog's developments in maximum likelihood factor analysis (1966, 1967, 1969). While Jöreskog's methods are derived in the statistics of confirming hypotheses, usage of alternate hypotheses brings this approach back toward exploratory methodology. Material in chapters 7 through 13 discusses methods from the exploratory side of the continuum.

A philosophic point presented in preceding paragraphs that almost all acceptable factor analytic models will <u>not</u> fit the real world needs to be emphasized for exploratory factor analytic studies. For experimenters and analysts to insist on a perfect fit in the population at all times is a route to chaos. For some factor methods a statistic has been developed to test whether analysis results fit the world in the population. Findings with large samples have been that quite reasonable analysis results have been rejected by these statistical tests. This point was first indicated by Tucker and Lewis (1973) who proposed a reliability coefficient for maximum likelihood factor analysis. Tucker and Lewis proposed this coefficient as a measure of goodness of fit. Many practitioners have come to realize the point that they should not expect acceptable models to fit perfectly in the population and have converted the chi square statistic to a measure of goodness of fit. A variety of other measures of goodness of fit have been proposed. One feature being investigated is the relation of these measures of goodness of fit to sample size, an ideal being that a measure of goodness of fit should not be affected by sample size on the average. An acceptable measure should estimate how well an analysis results fit in the population and should not be biased by sample size. The chi square statistic is biased so as to favor small samples. To be sure, however, large enough batteries of attribute measures and samples of individuals should be used to yeild stable analysis results; but, there should not be an insistence that the factor analysis model fit precisely.

A related problem concerns relevant small effects which might be missed when results of an analysis are accepted with some extent of lack of fit. To be sure, we should avoid missing small factors which can be made large and important with special studies. All experimenters should be alert to the possibility that there are small, but important factors in their data. They should inspect residuals after extraction of a factor analytic model from a body of data for indications of possibly important factors that have been missed in the design of the study and which could be strengthend in future studies. This distinction between trivia and important small influences is a matter for experimenter insight and judgement. However, the complexities of factor analysis make it imperative that such a distinction be made. Inclusion of a number of very small factors in a study results in an unmanageably large dimensionality of the common factor space. Great care is required of an experimenter in making the decisions between possibly meaningful factors and trivia. There should be no doubt, however, that some trivia will exist so that an acceptable factor analytic model will not fit the data perfectly, nor would the model fit perfectly in the population. Experimenters should be prepared, on the other hand, for the possibility that no acceptable factor model can fit their data. That the factor analytic methodology does not fit all bodies of data is not a good reason to discard factor analytic methodology. Many factor analytic studies yield valuable results even though the fits to the data are imperfect.

7.1. Review of Basic Factor Analytic Formulas

Basic formulas presented in Chapters 3 and 4 are repeated here for convenience of reference. A minimum of description will be given here since more extensive discussions were given in the preceding chapters.

Relations among score vectors are given in equation (3.1).

$$\boldsymbol{y} = \underline{z} + \ddot{\underline{\boldsymbol{z}}} \tag{3.1}$$

where $\boldsymbol{y}$ is the score vector on observed attributes for an individual, $\underline{z}$ is the modeled score vector for the individual from the factor analytic model, and $\ddot{\underline{\boldsymbol{z}}}$ is the vector of discrepancies of fit of the factor model to the observed scores.

Covariance relations in the population are given by the following formulas.

$$\boldsymbol{\Sigma}_{yy} = \boldsymbol{\Sigma}_{zz} + \boldsymbol{\Sigma}_{z\ddot{z}} + \boldsymbol{\Sigma}_{\ddot{z}z} + \boldsymbol{\Sigma}_{\ddot{z}\ddot{z}} \tag{3.88}$$

$$\boldsymbol{\Delta}_{\Sigma} = \boldsymbol{\Sigma}_{z\ddot{z}} + \boldsymbol{\Sigma}_{\ddot{z}z} + \boldsymbol{\Sigma}_{\ddot{z}\ddot{z}} \tag{3.89}$$

$$\boldsymbol{\Sigma}_{yy} = \boldsymbol{\Sigma}_{zz} + \boldsymbol{\Delta}_{\Sigma} \tag{3.90}$$

$\boldsymbol{\Sigma}_{yy}$ is the covariance matrix among the observed attributes, $\boldsymbol{\Sigma}_{zz}$ is the covariance matrix among the modeled attributes, $\boldsymbol{\Delta}_{\Sigma}$ is the discrepancy of fit of the model covariance matrix to the covariance matrix among the observed attributes in the population. Properties of $\boldsymbol{\Delta}_{\Sigma}$ will be discussed in subsequent paragraphs.

The factor model for covariance relations in the population is given by the following formulas.

$$\boldsymbol{\Sigma}_{zz} = \boldsymbol{B\Phi B}' + \boldsymbol{U}^2 \tag{3.35}$$

where $\boldsymbol{B}$ is the matrix of weights on the common factors, $\boldsymbol{\Phi}$ is the matrix of intercorrelations among the common factors, $\boldsymbol{U}^2$ is the diagonal matrix of unique variances. Score vectors on common factors and unique factors are designated by $\underline{\boldsymbol{x}}_{\beta}$ and $\underline{\boldsymbol{x}}_{\mu}$ with the modeled score vector being given by the following formula.

$$\underline{z} = \underline{\boldsymbol{x}}_{\beta}\boldsymbol{B}' + \underline{\boldsymbol{x}}_{\mu}\boldsymbol{U}' \tag{3.39}$$

A trait matrix, $\boldsymbol{T}$, is defined by the following formula.

$$\boldsymbol{\Phi} = \boldsymbol{TT}' \tag{3.42}$$

A matrix of factor weights, $\boldsymbol{A}$, for uncorrelated factors is defined by the following formula

$$\boldsymbol{A} = \boldsymbol{BT} \tag{3.44}$$

The factor model for uncorrelated common factors in the population is given below.

$$\boldsymbol{\Sigma}_{zz} = \boldsymbol{AA}' + \boldsymbol{U}^2 \tag{3.45}$$

Covariance relations in a sample are given by the following formulas.

$$\boldsymbol{C}_{yy} = \boldsymbol{C}_{zz} + \boldsymbol{C}_{z\ddot{z}} + \boldsymbol{C}_{\ddot{z}z} + \boldsymbol{C}_{\ddot{z}\ddot{z}} \tag{4.22}$$

$$\boldsymbol{\Delta}_{c} = \boldsymbol{C}_{z\ddot{z}} + \boldsymbol{C}_{\ddot{z}z} + \boldsymbol{C}_{\ddot{z}\ddot{z}} \tag{4.23}$$

$$\boldsymbol{C}_{yy} = \boldsymbol{C}_{zz} + \boldsymbol{\Delta}_{c} \tag{4.24}$$

$\boldsymbol{C}_{yy}$ is the covariance matrix among the observed attributes, $\boldsymbol{C}_{zz}$ is the covariance matrix among the modeled attributes, $\boldsymbol{\Delta}_{c}$ is the discrepancy of fit of the model covariance matrix to the covariance matrix among the observed attributes in the sample. Properties of $\boldsymbol{\Delta}_{c}$ will be discussed in subsequent paragraphs.

In a sample, the factor model fitted to an observed covariance matrix, $\boldsymbol{C_{yy}}$, is given by equation (4.29). Note that the estimated factor weight matrix, $\boldsymbol{B}$, is for common factor scores standardized in the sample. Matrix $\boldsymbol{R_{bb}}$ (note change in subscripts from $\beta\beta$ to bb) contains estimated intercorrelations among common factors and matrix $\boldsymbol{U^2}$ contains estimated unique variances.

$$\boldsymbol{C^{+}_{zz} = BR_{bb}B' + U^2} \tag{4.29}$$

$$\boldsymbol{C_{yy} = C^{+}_{zz} + \Delta^{+}_{y}} \tag{4.30}$$

The following formulas for operations in a sample parallel formulas given for transformations to uncorrelated factors in the population.

$$\boldsymbol{R_{bb} = TT'} \qquad \text{(7.1, parallel to 3.42)}$$

$$\boldsymbol{A = BT} \qquad \text{(7.2, parallel to 3.44)}$$

$$\boldsymbol{C^{+}_{zz} = AA' + U^2} \qquad \text{(7.3, parallel to 3.45)}$$

Matrix $\boldsymbol{T}$ is the trait matrix with trait vectors as rows. Matrix $\boldsymbol{A}$ contains the factor loadings on uncorrelated factors.

A geometric representation in the common factor space was presented in Chapter 3. Parallel equations for a sample are given here. In addition to the trait matrix $\boldsymbol{T}$ and the uncorrelated factor matrix $\boldsymbol{A}$, four new matrices are defined. Matrix $\boldsymbol{Q}$ contains covariances of modeled attributes with traits; matrix $\boldsymbol{F}$ contains, as rows, normals to hyperplanes; matrix $\boldsymbol{G}$ contains projections of the modeled attributes on the normals, these being part correlations of the modeled attributes with the factors other factors being partialled out; and matrix $\boldsymbol{D}$ contains cosines of angles between the trait vectors and the normals to the hyperplanes.

$$\boldsymbol{B = AT^{-1}} \qquad \text{(7.4, parallel to 3.48)}$$

$$\boldsymbol{Q = AT'} \qquad \text{(7.5, parallel to 3.55)}$$

$$\boldsymbol{TF' = D} \qquad \text{(7.6, parallel to 3.56)}$$

$$\boldsymbol{F' = T^{-1}D} \qquad \text{(7.7, parallel to 3.58)}$$

$$\boldsymbol{D = (Diag(R^{-1}_{bb}))^{-\frac{1}{2}}} \qquad \text{(7.8, parallel to 3.60)}$$

$$\boldsymbol{G = AF'} \qquad \text{(7.9, parallel to 3.61)}$$

$$\boldsymbol{G} = \boldsymbol{B}\boldsymbol{D} \qquad (7.10, \text{parallel to } 3.63)$$

7.2. Comments on Fitting a Factor Model to a Covariance Matrix

Almost always the distinctive portion of a factor analytic study starts with a covariance matrix. It is important, however, to remember that a factor analytic study starts with the design of the study, continues with selection and construction of attribute measures, arrangements for a sample of individuals, administration of the attribute measures to the individuals in the sample, scoring of the attribute measures and computation of the covariance matrix. These matters have been discussed in preceding chapters. Note that the analytic portion distinctive to factor analysis does not start with the attribute measures vectors, $\underline{\boldsymbol{y}}$, for individuals in the sample. The reason for this is a matter related to factor score indeterminacy, a topic to be discussed in score indeterminacy, a topic to be discussed in a later chapter. Only the nature of the source of the problems in this area will be indicated here.

As given in equation (3.39), modeled score vector, $\underline{z}$, can be expressed in terms of a contribution from the common factors and a contribution from the unique factors. Let $\underline{\boldsymbol{z}}_{\beta}$ and $\underline{\boldsymbol{z}}_{\mu}$ symbolize these two contribution vectors:

$$\underline{\boldsymbol{z}}_{\beta} = \underline{\boldsymbol{x}}_{\beta}\boldsymbol{B}' \qquad (7.11)$$

$$\underline{\boldsymbol{z}}_{\mu} = \underline{\boldsymbol{x}}_{\mu}\boldsymbol{U}' \qquad (7.12)$$

Then, from equation (3.39), the modeled score vector can be expressed as:

$$\underline{z} = \underline{\boldsymbol{z}}_{\beta} + \underline{\boldsymbol{z}}_{\mu} \qquad (7.13)$$

or

$$\underline{\boldsymbol{z}}_{\beta} = \underline{z} - \underline{\boldsymbol{z}}_{\mu} \qquad (7.14)$$

Combining these equations with equation (3.1) yields:

$$\underline{\boldsymbol{z}}_{\beta} = \underline{\boldsymbol{y}} - \underline{\boldsymbol{z}}_{\mu} - \underline{\ddot{\boldsymbol{z}}} \qquad (7.15)$$

The crux of the problem is that the score vectors $\underline{\boldsymbol{z}}_{\mu}$ and $\underline{\ddot{\boldsymbol{z}}}$ are not defined and until they are defined, $\underline{\boldsymbol{z}}_{\beta}$ can not be determined. If $\underline{\boldsymbol{z}}_{\beta}$ were determined, the solution in the common factor space would result from a principal components analysis as per equation (7.11). However, vector $\underline{\boldsymbol{y}}$ is in an *n* dimensional space (*n* being the number of measured attributes). The combined vector [$\underline{\boldsymbol{x}}_{\beta}$, $\underline{\boldsymbol{x}}_{\mu}$] is in an *r* + *n* dimensional space (*r* being the number of common factors). Vector $\underline{z}$ is in an *n* dimensional subspace of the combined vector [$\underline{\boldsymbol{x}}_{\beta}$, $\underline{\boldsymbol{x}}_{\mu}$]. Given a solution in the population to equations (3.59) and (3.90) with a satisfactory discrepancy matrix $\boldsymbol{\Delta}_{\Sigma}$, knowledge of $\boldsymbol{B}$, $\boldsymbol{\Phi}$, $\boldsymbol{U^2}$ and the modeled score vector $\underset{\sim}{z}$ yields not one but many solutions for score vectors $\underset{\sim}{\boldsymbol{x}}_{\beta}$ and $\underset{\sim}{\boldsymbol{x}}_{\mu}$ as shown by Guttman (1955). This is the factor score indeterminacy problem.

A further problem is introduced by a lack of knowledge of the discrepancies vectors $\underset{\sim}{\ddot{\boldsymbol{z}}}$. In fact, the nature of these vectors depends on the method of fitting the model to the observations of measured attributes. Each method for fitting a model to observations depends on a measure of goodness of fit, or of lack of fit; and, when different measures of lack of fit are tried, different solutions are obtained for parameters of the model. Consider a very simple example of this general principle: we have a collection of observations of the value of a variable and wish a single, representative value. The arithmetic mean minimizes the sum of squared deviations, this being the measure of lack of fit. In contrast, the median minimizes the sum of absolute values of deviations, this being the alternate measure of lack of fit. It is well known that the arithmetic mean and the median are not necessarily equal. One might consider any of a number of other possible lack of fit measures and arrive at a number of other representative values for our collection of observations. Fitting a factor model to observations both at the score level and at the covariance level possesses this same general problem, but in greater complexity. Most lack of fit measures, or loss functions used in factor analysis are stated in terms of covariance or correlation matrices. Some are more explicit than others, for example, the loss function for matrix factoring is more obscure than the loss function for Jöreskog's and Goldberger's generalized least squares factor analysis (1972). However, the application of these loss functions at the score level are not well known. Consequently, the score discrepancy vectors $\underset{\sim}{\ddot{\boldsymbol{z}}}$ are not well defined.

As a consequence of the indeterminacies of score vectors $\underset{\sim}{\boldsymbol{z}}_{\beta}$, $\underset{\sim}{\boldsymbol{z}}_{\mu}$ and $\underset{\sim}{\ddot{\boldsymbol{z}}}$, factor analysis theory has been extended analytically to derived statements for covariance and correlation matrices. While factor scores can not be determined uniquely, analysis of the covariance and correlation structures will yield important information about the structure of behavior. Estimates of the factor scores are available for use in particular situations; the estimation procedures and problems will be discussed in a subsequent chapter. Considerations of covariance matrices in the population led to equations (3.35) and (3.90) and in a sample to equations (4.29) and (4.30). Further considerations of the transformation of factors problem led to alternate equations for uncorrelated factors, equation (3.45) in the population and equation (7.3) in a sample. Substitution from equation (7.3) into (4.30) yields:

$$\boldsymbol{C_{yy}} = \boldsymbol{AA'} + \boldsymbol{U^2} + \boldsymbol{\Delta_y^+} \tag{7.16}$$

This is the formula most frequently used in practice in fitting a factor analytic model to observations. Matrix $\boldsymbol{\Delta_y^+}$ often is termed the matrix of residual covariances (or residual correlations in case $\boldsymbol{C_{yy}}$ is a correlation matrix) or, for short, the matrix of residuals. A number of the methods for fitting a model to an observed covariance matrix involve a measure of lack of fit, or loss function derived from matrix $\boldsymbol{\Delta}_y^+$. Further discussion of problems in measures of lack of fit, or loss functions will be presented in chapter 9.

A serious problem not dealt with previously in this chapter is the number of common factors to use in the fitted model. This is the number of columns in matrix $\boldsymbol{A}$ and the dimensionality of the common factor space. Several suggestions have been made as a result of serious study and will be described in subsequent chapters. Further, several suggestions have been made for supplementary information relative to this problem. Frequently, for any given method of fitting the factor model, solutions are obtained for several successive numbers of factors and the resulting measures of goodness of fit are compared. A number of measures of lack of fit (for which smaller values are better) reduce as the number of factors is increased. The question becomes whether the reduction in lack of fit from one number of factors to a larger number is warranted considering the additional number of entries in matrix $\boldsymbol{A}$. Detailed consideration of the number of factors will be given in following chapters; it is sufficient, here, to recognize this problem and to keep it in mind during following discussions.

We turn now to several older factor extraction methods which were applied to correlation matrices. Equation (7.16) may be rewritten in terms of the correlation matrix, $\boldsymbol{R}_{yy}$, among measured attributes.

$$\boldsymbol{R}_{yy} = \boldsymbol{A}\boldsymbol{A}' + \boldsymbol{U}^2 + \boldsymbol{\Delta}_R \tag{7.17}$$

where $\boldsymbol{\Delta}_R$ is the matrix of discrepancies in fitting the model for the correlation matrix. Equation (7.17) may be rearranged to:

$$(\boldsymbol{R}_{yy} - \boldsymbol{U}^2) = \boldsymbol{A}\boldsymbol{A}' + \boldsymbol{\Delta}_R \tag{7.18}$$

In this form, factor extraction involved establishing some guess at the unique variances in $\boldsymbol{U}^2$ so as to establish the matrix of correlations with communalities in the diagonal, this matrix being $(\boldsymbol{R}_{yy} - \boldsymbol{U}^2)$. Having this matrix, the procedures turned to factor extraction. We, knowingly, use the terms guessed uniqueness and guessed communalities since precise values do not exist in the population as a result of the idea that they are not defined until a method of fitting the factor analytic model has been selected. Two methods for obtaining guessed communalities will be presented in the next chapter. Also to be presented is Guttman's general theory for matrix factoring (1944) which gives a mathematical basis of factor extraction. Mehtods of factor extraction to be considered include the centroid method developed by L. L. Thurstone (1935, 1947), the group centroid method developed in several forms by Thurstone (1945) and discussed by Guttman (1952). An initial discussion will be presented in chapter 8 of principal factors extraction with a more extended discussion given in chapter 9 which includes determination of cormnunalities by a least squares fitting procedure. Principal factor extraction from a correlation matrix having guessed communalities is a commonly used and highly recommended method for exploratory factor analysis.

Detailed discussion of factor extraction by matrix factoring techniques appear in chapter 8. Chapter 9 will cover factor extraction by methods using loss functions. Most of these techniques result in a factor matrix for uncorrelated factors which require transformation.

7.3. Properties of Transformations of Factors

A most important problem for exploratory factor analysis follows establishment of a common factor matrix $\boldsymbol{A}$ with uncorrelated factors. As discussed in Chapter 3 in terms of a geometric representation there is almost complete freedom in establishing trait vectors in matrix $\boldsymbol{T}$ and corresponding loadings on the derived correlated factors. Basic equations for operations in a sample are given in equations (7.4) through (7.10). Only two restrictions exist for matrix $\boldsymbol{T}$ in maintaining the representation of the input covariance or correlation matrix.
First: matrix $\boldsymbol{T}$ must be square and nonsingular so that an inverse exists:

$$|\boldsymbol{T}| \neq \boldsymbol{0} \qquad (7.19)$$

Second: each trait vector (row of matrix $\boldsymbol{T}$) must be of unit length so that:

$$Diag(\boldsymbol{TT'}) = Diag(R_{bb}) = \boldsymbol{I} \qquad (7.20)$$

Many matrices $\boldsymbol{T}$ may be written which satisfy the preceding conditions. Consider matrix $\boldsymbol{C_c}$ to be the portion of $\boldsymbol{C^{+}_{zz}}$ from the common factors. $\boldsymbol{C_c}$ is the covariance matrix among the score vectors $\underline{\boldsymbol{z}}_{\boldsymbol{\beta}}$. Then, from equation (4.29) and (7.3):

$$\boldsymbol{C_c} = \boldsymbol{BR_{bb}B'} = \boldsymbol{AA'} \qquad (7.21)$$

and

$$\boldsymbol{C^{+}_{zz}} = \boldsymbol{C_c} + \boldsymbol{U^2} \qquad (7.22)$$

There is an urgent need for some additional principle for use in developing a useful transformation of factors.

Thurstone (1935, 1947) proposed the principle of simple structure to resolve the problem of definition of the transformation of factors for many areas of study. His reasoning was that, for a diverse battery of attributes, each factor should not influence the scores on all attributes. This line of reasoning leads to a conclusion that there should be a number of zero or trivial loadings on each transformed factor. Transformation of factors would be a search for a location of trait vectors so as to have a number of zero or trivial loadings. This principle is widely accepted and has led to inclusion in most computer programs of analytic procedures which tend toward a simple structure solution. Maybe, this acceptance has been overly wide spread and has led in many cases to uncritical acceptance of computer output. Every analyst has a responsibility to inspect results from a study to judge whether or not the principle of simple structure applies to the data at hand. Other principles may exist for other classes of cases than that for which simple structure was suggested. Maybe, of course, a given body of data is inadequate to define a simple structure. A quite serious possibility is that the computerized solution does not yield satisfactory results even when an acceptable simple structure exists for a given factor matrix $\boldsymbol{A}$. However, simple structure appears to apply to many bodies of data so as to yield useful conjectures as to the nature of constructs and dynamics underlying the observed covariations among the surface attributes.

The word structure is used in terms of the geometric representation of the common factor space. It includes the configuration of attribute vectors, a collection of trait vectors and hyperplanes along with the normals to these hyperplanes. In the sense used here, a structure is analogous to a building which includes all of the material in place in the building including the structural steel. A structure includes all of the parts and their relations to each other in an object. In the case of the common factor space, the parts are elements of the geometric representation.

For an example of a common factor vector configuration consider the covariance matrix $\boldsymbol{C_c}$ given in Table 7.1. Diagonal entries are the squares of the lengths of the vectors, the vector lengths being labeled h_j (square roots of the diagonal entries). Off diagonal entries are scalar products between pairs of vectors. Let θ_{jk} designate the angle between vectors j and k ; then:

$$C_{cjk} = \mathrm{h}_j \mathrm{h}_k cos\theta_{jk} \tag{7.23}$$

Conversion of the covariance matrix $\boldsymbol{C_c}$ in Table 7.1 to a correlation matrix R_c of Table 7.2 involves division of each covariance by the product of the square roots of the corresponding diagonal entries. Thus:

$$r_{cjk} = C_{cjk}/(\mathrm{h}_j \mathrm{h}_k) \tag{7.24}$$

which, with equation (7.23), yields:

$$r_{cjk} = cos\theta_{jk} \tag{7.25}$$

Given the correlations in Table 7.2, the angles between the vectors may be obtained from a table of trigonometric functions, the angles for the artificial example being given in Table 7.3. Two views of a vector configuration for our example are given in Figure 7.1.

A configuration of attribute vectors may be constructed from a covariance matrix by computing the vector lengths and angles between pairs of vectors as described in the preceding

Table 7.1

Covariance Matrix from Common Factors among Attributes

	1	2	3	4	h_j
1	.25				.5000
2	.03	.34			.5831
3	.22	.27	.37		.6083
4	.07	.31	.28	.29	.5385

Table 7.2

Correlation Matrix from Common Factors among Attributes

	1	2	3	4
1	1.00			
2	.10	1.00		
3	.72	.76	1.00	
4	.26	.99	.85	1.00

Table 7.3

Angles (in Degree) between Attribute Vector

	1	2	3	4
1	--			
2	84.09	--		
3	43.67	40.43	--	
4	74.93	9.16	31.26	--

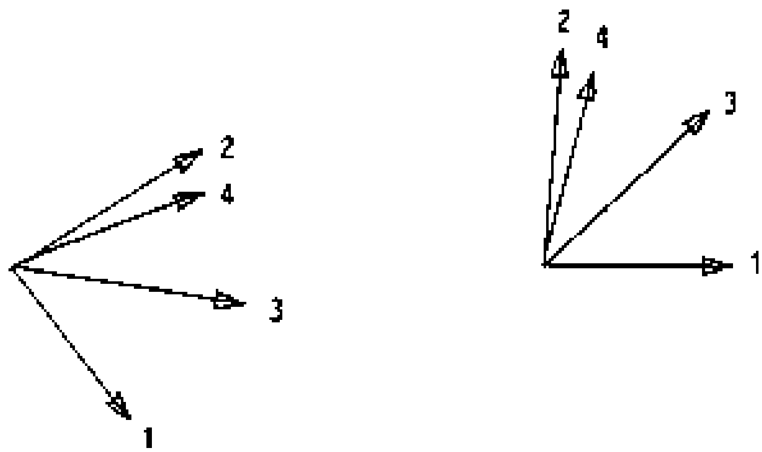

Figure 7.1: Two Views of a Geometric Vector Configuration
Representing Common Factor Covariance Matrix

paragraph and following the procedure to be described. This procedure does not depend upon obtaining coordinates of the terminals of the vectors on coordinate axes. Lay out vector 1 in any direction with length h_1. Lay out vector 2 at an angle θ_{12} with vector 1 and of length h_2. Best results are obtained by choosing vectors 1 and 2 to have an angle as near to 90° as possible. To achieve this result the attributes may have to be reordered. In the example no reordering was necessary. A third vector is laid out in a direction with angles θ_{13} and θ_{23} from the first two vectors and with length h_3. Some times a third dimension will be required; however, in no case will it be impossible to lay out this third vector. In some cases, as in the example, the third vector will lie in the plane defined by the first two vectors. Vector 4 is laid out in a similar fashion in terms of angles θ_{14} **,** θ_{24} and θ_{34} with length h_4. Additional attribute vectors could be added to the configuration in terms of the vectors' angles with preceding vectors and with the given lengths. With the addition of each successive vector a new dimension may have to be utilized; however, as remarked earlier, in no case will it be impossible to add a new attribute vector. This condition is normal for a covariance matrix so that when it is not possible to add a new vector there is a blunder in the covariance matrix. The number of dimensions required equals the rank of the covariance matrix. In our example the rank of the covariance matrix, $\boldsymbol{C_c}$ is 2 which equals the number of dimensions of the vector configuration.

Figure 7.2 presents a supplementary view of the vector configuration. All attribute vectors have been lengthened to unit length. Then, the relations among the vectors are determined by the correlation matrix, $\boldsymbol{R_c}$, by the angles between the pairs of vectors. Once the unit length vectors have been established, the original length vectors are established with the original vector lengths.

A most important point is that the vector configuration is determined solely by the covariance matrix $\boldsymbol{C_c}$. No axes have been defined. Given the covariance matrix $\boldsymbol{C_c}$, different methods of factor extraction will insert axes and yield matrices of coordinates on the axes. The locations of these axes will be different for different methods of factor extraction. Two pairs of orthogonal axes have been inserted in Figure 7.3. The first pair is labeled I and II while the second pair is labeled P and Q . Coordinates of the terminals of the attribute vectors on these pairs of axes are given in Table 7.4 as factor matrices $\boldsymbol{A_1}$ and $\boldsymbol{A_2}$. These factor matrices satisfy the condition that:

$$\boldsymbol{C_c} = \boldsymbol{A_1 A_1'} = \boldsymbol{A_2 A_2'} \tag{7.26}$$

Axis I was passed through the centroid of the attribute vectors, the centroid vector equaling the sum of the attribute vectors divided by the number of attributes. Axis II was set orthogonal to axis I . This is the result of the centroid method of factor extraction. For the second pair of axes, axis P was passed through the vector for attribute 1 and axis Q was set orthogonal. This is the result of the diagonal method of factor analysis which is related to the Cholesky decomposition

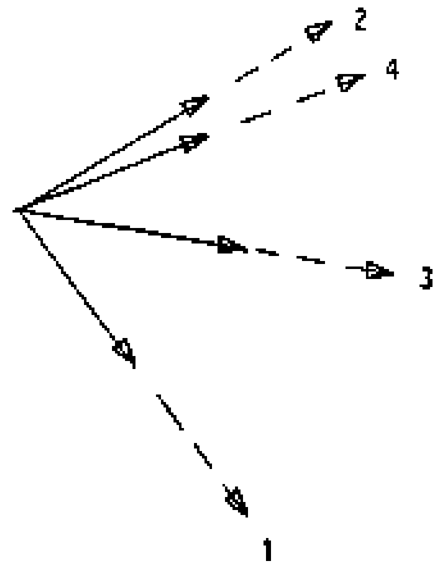

Figure 7.2: Construction of a Geometric Vector Configuration from
Common Factor Covariance and Correlation Matrices

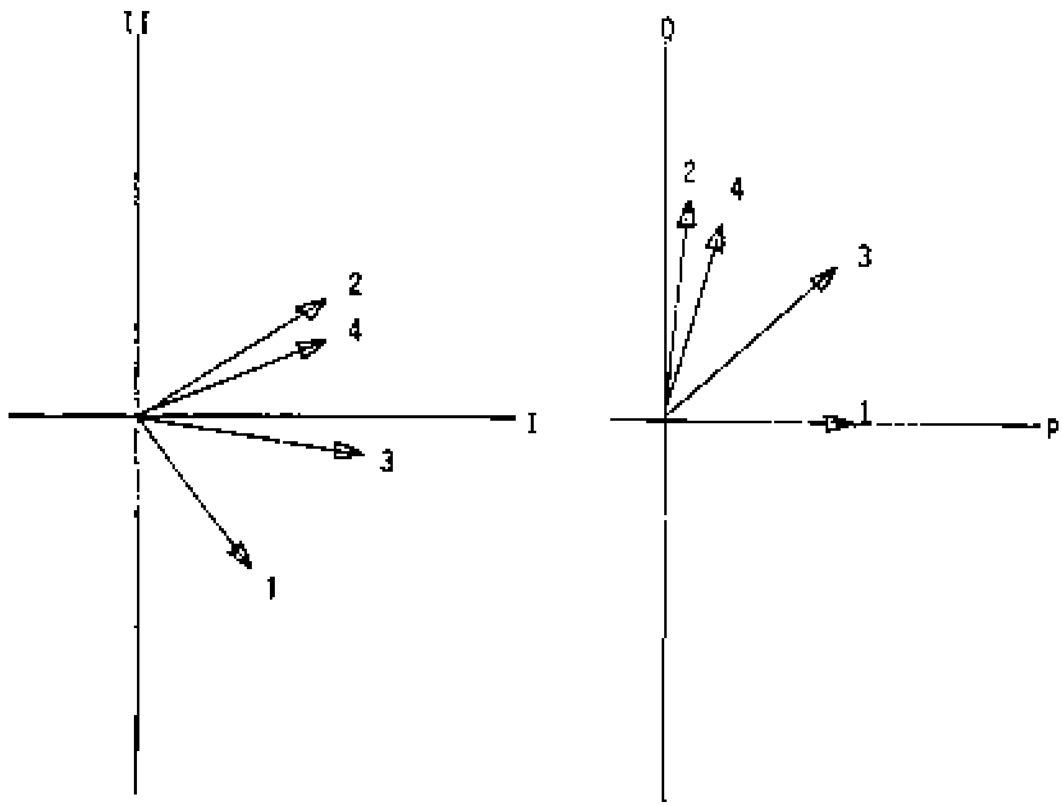

Figure 7.3: Geometric Vector Configuration with Two Sets of Coordinate Axes

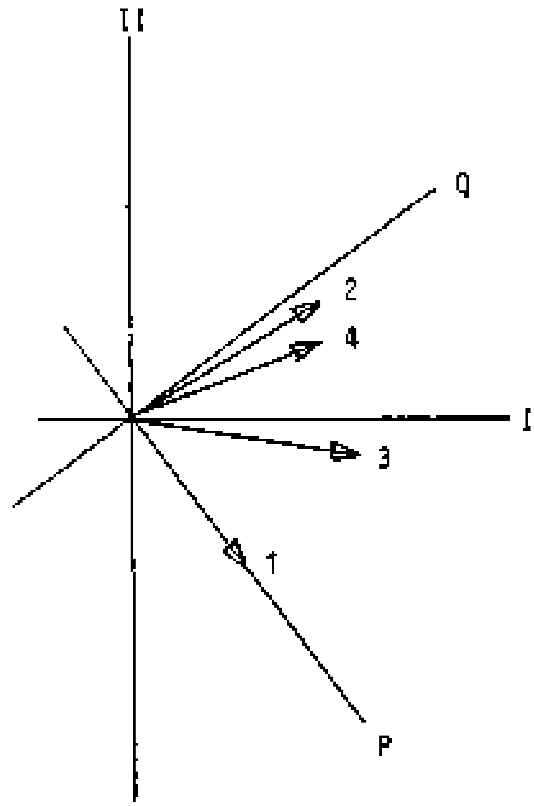

Figure 7.4: Transformation from One Set of Coordinate Axes to a Second Set

Table 7.4

Two Uncorrelated Factor Matrices from C_c

	A_1			A_2	
	I	II		P	Q
1	.30	-.40	1	.50	.00
2	.50	.30	2	.06	.58
3	.60	-.10	3	.44	.42
4	.50	.20	4	.14	.52

Table 7.5

Transformation Matrix from Factor Matrix A_1 to Factor Matrix A_2

	I	II
P	.60	-.80
Q	.80	.60

Table 7.6

Normalized Attribute Vectors from factor Matrix A_1

	A_1				Normalized A_1	
	I	II	h_j		P	Q
1	.30	-.40	.5000	1	.60	-.80
2	.50	.30	.5831	2	.86	.51
3	.60	-.10	.6083	3	.99	-.16
4	.50	.20	.5385	4	.93	.37

of a symmetric matrix. Both the centroid method and the diagonal method of factor extraction will be described in the next chapter.

Figure 7.4 shows the attribute vector configuration with both pairs of axes and Table 7.5 presents the relation between the two pairs of axes. Let $\boldsymbol{M_{12}}$ be a square, orthonormal matrix. Then:

$$\boldsymbol{M}_{12} = \boldsymbol{M}_{12}^{-1} \tag{7.27}$$

so that

$$\boldsymbol{M}_{12}\boldsymbol{M}'_{12} = \boldsymbol{M}_{12}\boldsymbol{M}_{12} = \boldsymbol{I} \tag{7.28}$$

Let $\boldsymbol{M_{12}}$ be determined so that:

$$\boldsymbol{A}_1\boldsymbol{M}'_{12} = \boldsymbol{A_2} \tag{7.29}$$

This is always possible whenever $\boldsymbol{A_1}$ and $\boldsymbol{A_2}$ have the same number of columns which equals the rank of $\boldsymbol{C_c}$ **.** The reverse transformation is:

$$\boldsymbol{A_2}\boldsymbol{M}'_{12} = \boldsymbol{A}_1 \tag{7.30}$$

Table 7.5 gives the matrix $\boldsymbol{M_{12}}$ for our example. Row P describes axis P in terms of axes I and II while row Q describes axis Q in terms of axes I and II . Entries in $\boldsymbol{M_{12}}$ are the cosines of the angles of axes P and Q with axes I and II ; they are termed direction cosines. Each row is a unit length vector in the direction of the transformed axes, P and Q .

Unit length attribute vectors may be obtained by "normalizing" the vectors in an orthogonal factor matrix such as $\boldsymbol{A}_1$. The term "normalize" refers to adjusting the length of a vector to a unit length by dividing each coordinate of the vector by the length of the vector. Consider Table 7.6. Matrix $\boldsymbol{A_1}$ is given at the left for our artificial example. Column h_j contains the lengths of the attribute vectors. Since matrix $\boldsymbol{A_1}$ is an orthogonal factor matrix, being for uncorrelated factors, the squares of the vector lengths are the sums of the squares of the coordinates, or loadings, with the vector lengths being the square roots of these squares of vector lengths. The matrix "NORMALIZED $\boldsymbol{A}_1$" is given on the right of Table 7.6 with each row being the row of $\boldsymbol{A_1}$ divided by the h_j for that row. These normalized attribute vectors are plotted in Figure 7.5 with axes I and II . Since these normalized vectors are of unit length their termini are on a unit radius circle which has been drawn in Figure 7.5. This property of normalized vectors will be generalized for three dimensions to the termini laying on the surface of a unit radius sphere. For cases with more than three dimensions, the termini will lie on the surface of a unit hypersphere.

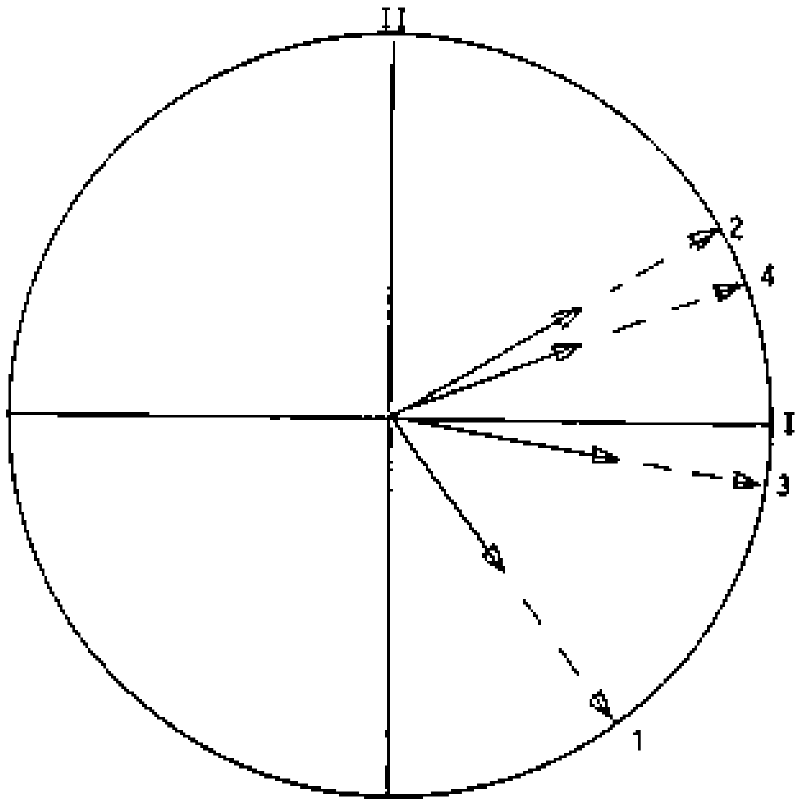

Figure 7.5: Normalized Attribute Vectors with a Set of Coordinate Axes

A drawing of a group of attribute vectors in three dimensions is given in Figure 7.6. This is a picture of a cardboard model on a stand with the vectors being drawn on three pieces of cardboard which form planes. These vectors are normalized, being of unit length, so that their termini would lie on the surface of a unit sphere. Figure 7.7 shows on the left three views of this model tipped forward by successive steps. The top view shows the outside of the front plane while the middle view is in line with this front plane and the bottom view shows the inside of the front plane. In the bottom view we are looking down into the model. On the right of Figure 7.7 are three corresponding drawings of a unit sphere with the termini of the attribute vectors indicated by closed circles. The planes of the model are shown also. Use of unit spheres is a convenient method for showing configurations of attribute vectors. Remember that the vectors emanate from the origin which is at the center of the sphere.

We return to the Nine Mental Tests example used in Chapter 1. Table 7.7 gives the factor matrix $\boldsymbol{A}$ for uncorrelated factors, the vector lengths in column h_j and the Normalized $\boldsymbol{A}$ computed by the procedure described in preceding paragraphs. Figure 7.8 shows the termini of the attribute vectors on the surface of a unit radius sphere. These termini are indicated by closed circles. A transformation of factors was computed by a method named DAPPFR which will be described in Chapter 11. Table 7.8 gives the resulting trait vectors and normal vectors. The trait vectors are rows of matrix $\boldsymbol{T}$ while the normal vectors are rows of matrix $\boldsymbol{F}$. Remember that the trait vectors and normal vectors are of unit length so that their termini would lie on the surface of the unit sphere. Figure 7.8 shows the trait vectors as open circles at the intersection of planes. The normal vectors are shown by crosses. The planes pass near to attribute vectors, this being a property of a successful transformation of factors to simple structure. Each normal vector is orthogonal to one of the planes and, thus, is orthogonal to every vector in the plane. Normal vector $\underline{\mathbf{f}}_1$ is orthogonal to the plane that passes through trait vectors $\underline{\mathbf{t}}_2$ and $\underline{\mathbf{t}}_3$. Consequently, normal vector $\underline{\mathbf{f}}_1$ is orthogonal to trait vectors $\underline{\mathbf{t}}_2$ and $\underline{\mathbf{t}}_2$. Note that normal vector $\underline{\mathbf{f}}_1$ is not orthogonal to trait vector $\underline{\mathbf{t}}_1$. In a similar fashion, each normal vector is orthogonal to all trait vectors except the corresponding trait vector. A result is that each trait vector is orthogonal to all normal vectors except the corresponding normal vector. The planes may be termed "base planes", or, in higher dimensions, "base hyperplanes". Each base plane may be considered as defined by the normal vector to that plane.

Three types of meaningful coefficients were described in Chapter 3 in terms of the geometric representation of the common factor model. These are structure coefficients relating the modeled attribute vectors to the trait vectors and normals. While they are computed from the common factor matrix $\boldsymbol{A}$ which contains vector coordinates on a set of orthogonal coordinate axes, these coefficients are properties of the common factor structure and are independent of the coordinate axes. They are not altered by an orthogonal rotation of the coordinate axes. Consider

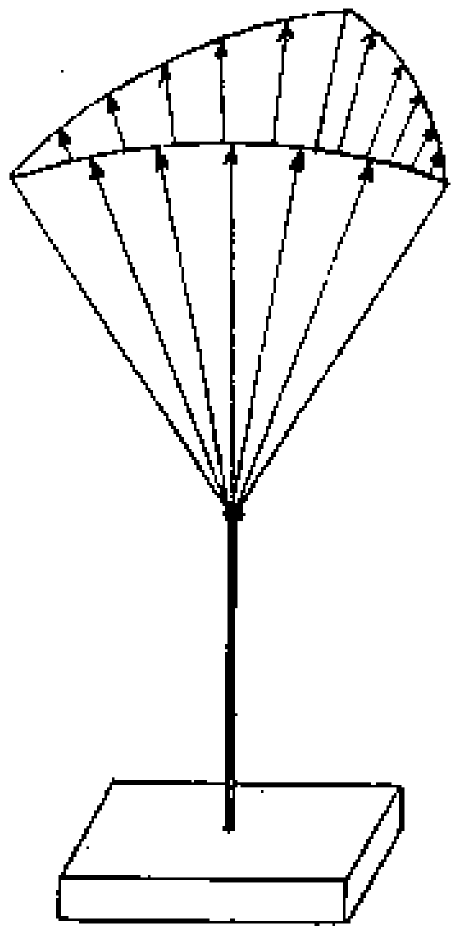

Figure 7.6: A Geometric Model of Normalized Vectors in Three Dimensions

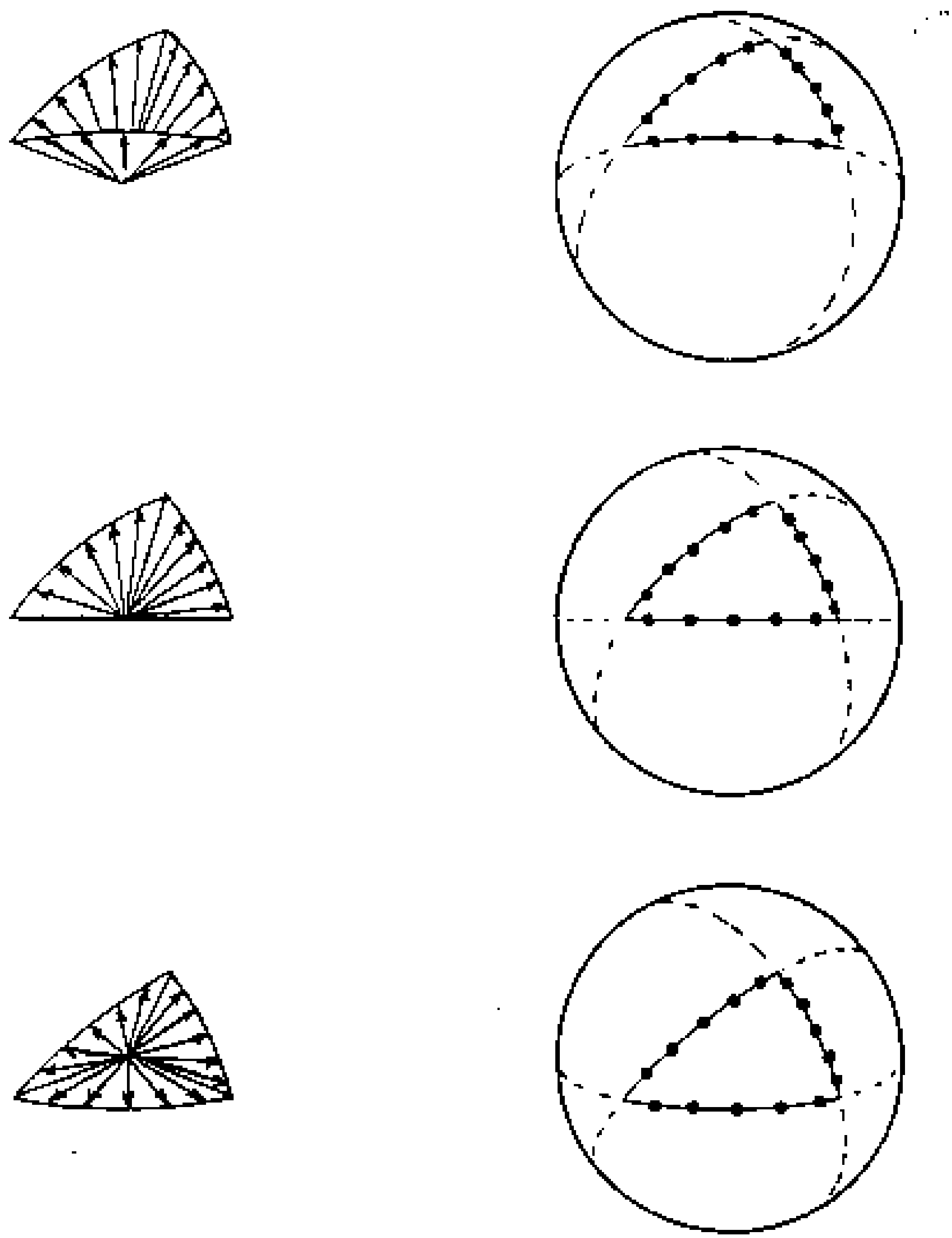

Figure 7.7: Three Views of the three Dimensional Vector Model with Rotation forward Accompanied by Spherical Representations of the Vector Terminals

Table 7.7

Uncorrelated Factor matrix for Nine Mental Tests

	Matrix A				Normalized A		
	1	2	3	h_j	1	2	3
1	.42	.36	.28	.6192	.67	.59	.45
2	.36	.54	.16	.7309	.64	.74	.21
3	.61	.16	.19	.6583	.92	.24	.30
4	.54	-.46	.09	.7126	.76	-.64	.13
5	.63	-.48	.05	.7960	.79	-.61	.06
6	.59	-.41	.06	.7161	.82	-.57	.09
7	.48	.40	-.21	.6618	.73	.61	-.31
8	.61	.00	-.29	.6738	.90	.00	-.44
9	.59	.16	-.25	.6603	.89	.24	-.38

Table 7.8

Factor Transformation Matrices for Nine Mental Tests

	Matrix T Trait Vectors			Matrix F Normal Vectors		
	1	2	3	1	2	3
1	.67	.67	.33	.31	.52	.80
2	.79	-.61	.09	.69	-.72	.08
3	.62	.53	-.58	.26	.20	-.94

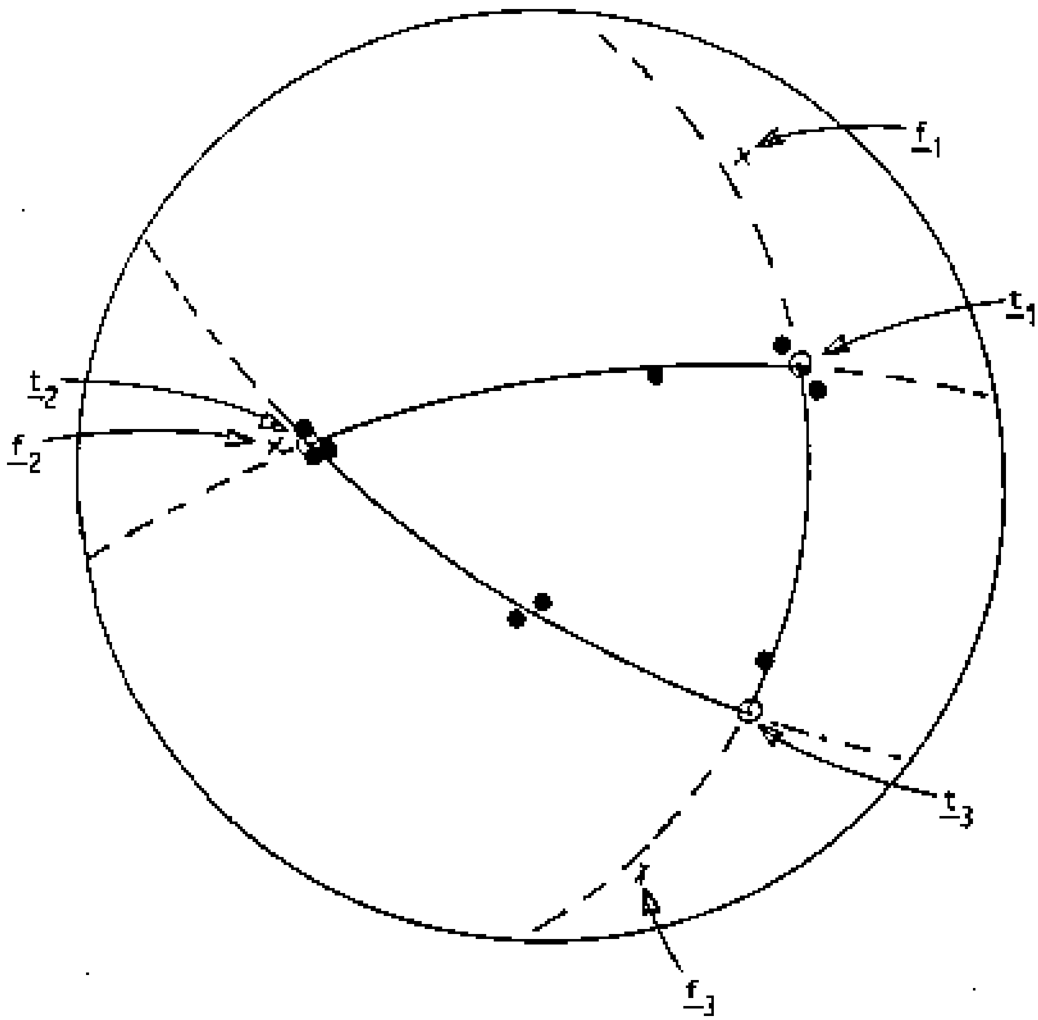

Figure 7.8: A Shperical Representation of the Nine Mental Tests Structure

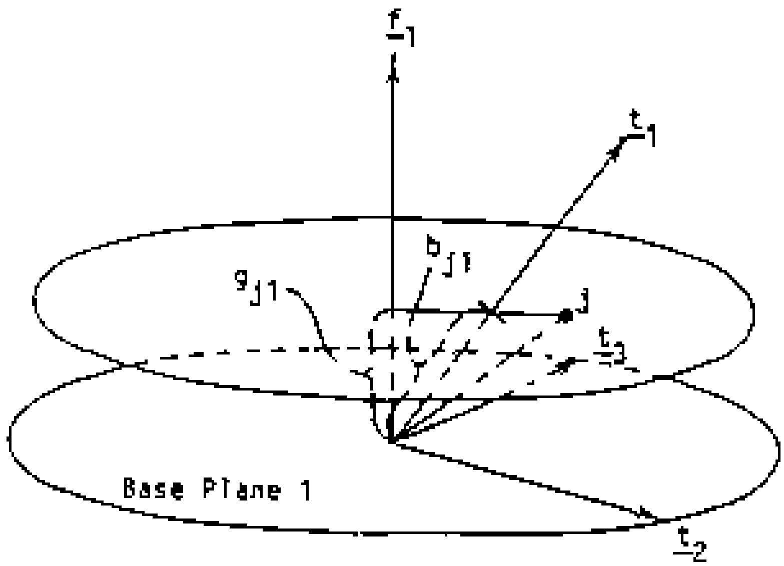

Figure 7.9: Illustration of Cartesian Coordinate and Projection on Normal in Three Dimensions

Table 7.9 which presents the matrices for the three types of structure coefficients for the nine mental test example. First is the matrix $\boldsymbol{B}$ of factor weights; second is the matrix $\boldsymbol{Q}$ of covariances of the modeled attributes with the factors; third is the matrix $\boldsymbol{G}$ of projections on the normals. Each of these will be discussed in turn.

As described in Chapter 3, the factor weights in matrix $\boldsymbol{B}$ are Cartesian coordinates of the attribute vectors on the trait vectors. In Figure 7.8 the base plane for factor 1 passes through trait vectors $\underline{\mathbf{t}}_2$ and $\underline{\mathbf{t}}_3$. This base plane is defined by normal $\underline{\mathbf{f}}_1$. Figure 7.9 gives an illustration of a Cartesian coordinate in three dimensions. Normal $\underline{\mathbf{f}}_1$ is vertical to base plane 1 horizontal. Trait vectors $\underline{\mathbf{t}}_2$ and $\underline{\mathbf{t}}_3$ are in this base plane. Attribute vector *j* is above the base plane and has had a projection plane passed through it parallel to the base plane. Trait vector $\underline{\mathbf{t}}_1$ passes through the projection plane at a distance b_{j1} from the origin. This is the Cartesian coordinate of attribute vector *j* with trait vector $\underline{\mathbf{t}}_1$. This is the geometric representation of a factor weight. It is the extent of the trait vector included in the attribute vector.

Structural coefficients in matrix $\boldsymbol{Q}$ are illustrated geometrically in Figure 7.10. These coefficients are the scalar products of the attribute vectors with the trait vectors. Since the trait vectors have unit length, the scalar products are the orthogonal projections of the attribute vectors on the trait vectors. Figure 7.10 is similar to Figure 7.9 involving the same normal $\underline{\mathbf{f}}_1$, base plane 1, trait vectors $\underline{\mathbf{t}}_1$, $\underline{\mathbf{t}}_2$ and $\underline{\mathbf{t}}_3$. The base plane 1 passes through trait vectors $\underline{\mathbf{t}}_2$ and $\underline{\mathbf{t}}_3$. Coefficient q_{jl} is the orthogonal projection of attribute vector *j* on trait vector $\underline{\mathbf{t}}_1$. This does not depend on the projection plane passed through the terminus of attribute vector *j* . As pictured, coefficient q_{jl} is larger than the factor weight b_{ji} shown in Figure 7.9. Note that the projection plane for an attribute vector in the base plane would have a zero factor weight while the orthogonal projection on the trait vector $\underline{\mathbf{t}}_1$ would not, necessarily, be zero. The coefficients in matrix $\boldsymbol{Q}$ are the covariances of the attributes with the traits.

Matrix $\boldsymbol{G}$ of projections on the normals is pictured in Figure 7.9 which shows the projection of attribute vector *j* on normal $\underline{\mathbf{f}}_1$. The projection plane parallel to the base plane and through the terminus of vector *j* cuts off a distance from the origin on normal $\underline{\mathbf{f}}_1$ equal to the projection g_{jl} . This projection is a measure of the contribution to attribute vector j which is independent of vectors in the base plane, including trait vectors $\underline{\mathbf{t}}_2$ and $\underline{\mathbf{t}}_3$. Statistically, these projections on the normals are semipartial covariances with trait vectors, having, for each trait vector, the effects of all other trait vectors partialled out.

As described in the preceding paragraphs, the structural coefficients in matrices $\boldsymbol{B}$, $\boldsymbol{Q}$ and $\boldsymbol{G}$ are geometrically independent of a reference set of axes. For mathematical and computational convenience the common factor structure is described in terms of an orthogonal axes frame starting with matrix $\boldsymbol{A}$. Matrices $\boldsymbol{T}$ of trait vectors and $\boldsymbol{F}$ or normals are referred to this same frame of orthogonal axes. The structural coefficients are independent of the location of

Table 7.9

Transformed Factor Matrices for Nine Mental Tests

(Structure Coefficient Matrices)

Matrix $B=AT^{-1}$

Factor Weights

	1	2	3
1 Addition	.66	.05	-.10
2 Multiplication	.67	-.05	.10
3 Three-Higher	.52	.32	.01
4 Figures	.00	.72	-.04
5 Cards	-.02	.79	.03
6 Flags	.02	.71	.02
7 Ident. Numbers	.23	.02	.49
8 Faces	-.06	.40	.54
9 Mirror Reading	.08	.28	.52

Matrix Q=AT'

Covariances with Factors

	1	2	3
1	.61	.13	.29
2	.72	.05	.48
3	.57	.40	.35
4	.08	.71	.04
5	.11	.80	.11
6	.14	.72	.11
7	.52	.11	.63
8	.31	.45	.55
9	.42	.34	.59

Matrix G=AF'

Projection on Normals

	1	2	3
1	.54	.05	-.08
2	.55	-.05	.08
3	.42	.32	.01
4	.00	.71	-.03
5	-.01	.79	.02
6	.02	.70	.02
7	.19	.02	.40
8	-.05	.39	.44
9	.06	.27	.42

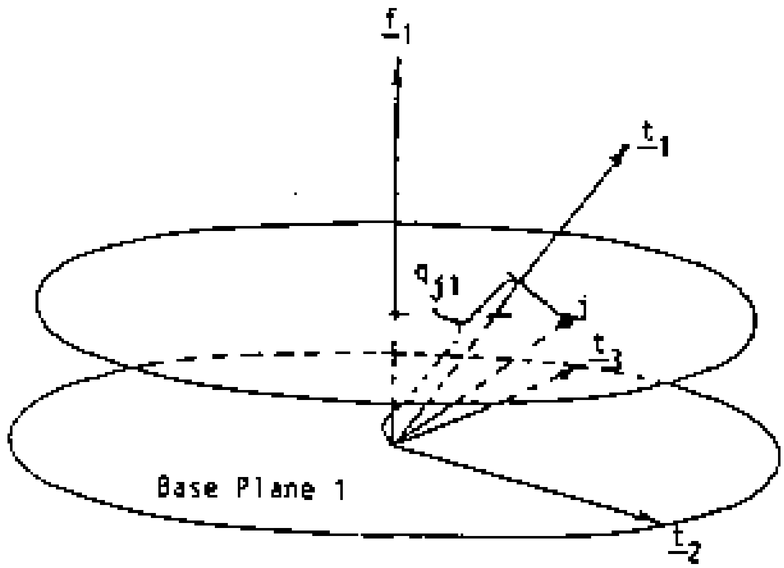

Figure 7.10: Illustration of Orthogonal Projection of an Attribute Vector
on a Trait Vector

the orthogonal axes. They are not altered by an orthogonal rotation of axes. Consider that there is one frame of orthogonal axes on which the vectors have coordinates in matrices $\boldsymbol{A}_1$, $\boldsymbol{T}_1$ and $\boldsymbol{F}_1$. Consider an orthogonal rotation of axes by a matrix $\boldsymbol{M}_{12}$ as was used in equation (7.29) to effect such a rotation.

$$\boldsymbol{A}_1\boldsymbol{M}'_{12} = \boldsymbol{A}_2 \tag{7.29}$$

This rotation can be carried to matrices $\boldsymbol{T}_1$ and $\boldsymbol{F}_1$:

$$\boldsymbol{T}_1\boldsymbol{M}'_{12} = \boldsymbol{T}_2 \tag{7.30}$$

$$\boldsymbol{F}_1\boldsymbol{M}'_{12} = \boldsymbol{F}_2 \tag{7.31}$$

From the orthogonal properties of matrix $\boldsymbol{M}_{12}$:

$$\boldsymbol{T}_2^{-1} = \boldsymbol{M}_{12}\boldsymbol{T}_1^{-1} \tag{7.32}$$

Then,

$$\boldsymbol{A}_2\boldsymbol{T}_2^{-1} = \boldsymbol{A}_1\boldsymbol{M}'_{12}\boldsymbol{M}_{12}\boldsymbol{T}_1^{-1} = \boldsymbol{A}_1\boldsymbol{T}_1^{-1} = \boldsymbol{B} \tag{7.33}$$

A similar demonstration of invariance is possible for matrices $\boldsymbol{Q}$ and $\boldsymbol{G}$.

Table 7.10 presents an outline of terminology used for the structural coefficients for attribute measures in the common factor space. We have presented three types of coefficients with their interpretations. A terminology has come into common use following Harman (1976). He pointed out the existence of a fourth type of coefficient in matrix $\boldsymbol{AF}^{-1}$. We do not have a ready interpretation for these coefficients. Computer packages frequently print matrices $\boldsymbol{B}$ and $\boldsymbol{G}$ with shortened titles of Pattern Loadings and Structure Loadings. We are using the term structure to refer to the collection of attribute vectors, trait vectors and normals to the hyperplanes along with the hyperplanes. Attachment of the term "structure loading" to one pair of types of coefficients appears to be a questionable interpretation of Thurstone's use of the word structure, as in simple structure. Thurstone used the projections on the normals as a measure of the independent contributions to the attributes.

Consider, now, the relations among and between the trait vectors and the normals. Table 7.11 presents matrices of these relations for the Nine Mental Tests example. Matrix $\boldsymbol{R}_{bb}$ has correlations among transformed factors, these correlations equaling the scalar products among the trait vectors. Since the trait vectors are unit vectors, these scalar products equal cosines of angles among the trait vectors. Matrix $\boldsymbol{FF}'$ contains scalar products (equal to cosines of angles) among the normals. Due to an inverse relation between $\boldsymbol{R}_{bb}$ and $\boldsymbol{FF}'$, the algebraic signs of the off-diagonal entries in these two matrices tend to be opposite. Matrix $\boldsymbol{D}$ contains scalar products between the trait vectors and the normals. Traits vectors and normals are paired,

Table 7.10
Outline of Structure Coefficients

Matrix	Terminology used here	Common terminology	
$AT^{-1} = B$	Factor Weights	Pattern loadings	on primary vectors
$AT' = Q$	Covariances with factors	Structure loadings	on primary vectors
AF^{-1}		Pattern loadings	on reference vectors
$AF' = G$	Projections on normals; semipartial covariances with factors	Structure loadings	on reference vectors

Table 7.11
Transformed Factors Relations Matrices for Nine Mental Tests

Matrix $R_{bb}=TT'$

	1	2	3
1	1.00	.14	.58
2	.14	1.00	.11
3	.58	.11	1.00

Matrix FF'

	1	2	3
1	1.00	-.10	-.57
2	-.10	1.00	-.03
3	-.57	-.03	1.00

Matrix D=TF'

	1	2	3
1	.81		
2		.99	
3			.82

one trait vector with one normal. The scalar product between the trait vector and the normal of a pair is a diagonal entry in $\boldsymbol{D}$. The off-diagonal entries in $\boldsymbol{D}$ are scalar products between non-paired trait vectors and normals. Each normal is orthogonal to all non-paired trait vectors. Likewise, each trait vector is orthogonal to all non-paired normals.

Various aspects of a structure in the common factor space have been discussed in the preceding paragraphs. Thurstone's conception of a simple structure is a particular form involving a theoretic postulate which may or may not be satisfied for any given body of data. As was presented earlier, the configuration of attribute vectors is defined by the covariance matrix $\boldsymbol{C_c}$ from the common factors. A structure is not complete until the trait vectors and normals have been defined. Thurstone's principal of simple structure involves a statement that the trait vectors could be located such that there would be many zero or near zero factor loadings. He argued that the nature of more basic factors should be such that each factor should influence scores on some attributes but not all attributes. If this be the case, trait vectors and normals could be located so that there were a number of trivial loadings on each factor. The hyperplanes could be defined by being near to a number of attribute vectors.

For examples of the concept of simple structure consider the nine mental tests example. In Figure 7.8 the base plane for factor 3 passes through trait vectors $\underline{\mathbf{t}}_1$ and $\underline{\mathbf{t}}_2$. This plane is moderately well defined by six points which are for the first six tests in the battery, see matrix $\boldsymbol{B}$ in Table 7.9, column 3 for factor 3. The last three tests have high loadings on this factor, they are the tests not near the base plane in Figure 7.8. These three tests involve highly speeded tasks of simple recognition of stimuli among distractors. This factor has been interpreted as a perceptual speed factor. While the first six tests were timed, performance depended more on conduct of mental tasks and to a trivial extent on speed of perceptual recognition of answers. By having only trivial dependence on perceptual speed, the first six tests determine the base plane for the third factor. The base plane for factor 1 passes through trait vectors $\underline{\mathbf{t}}_2$ and $\underline{\mathbf{t}}_3$. This plane is defined by five points in Figure 7.8 and is not as well determined as would be desirable. There is a point near trait vector $\underline{\mathbf{t}}_3$ near to the base plane but not "in" it. Consider the factor loadings for factor 1 in Table 7.9. The first three tests have high loadings; these tests involve numerical operations. Test 7, identical numbers, also involves simple numerical operations and has a lower loading on factor 1; it is the test with a vector near trait vector $\underline{\mathbf{t}}_3$. The base plane for factor 1 is defined by the vectors for the other five tests which do not involve numerical operations. The base plane for factor 2 is less well defined by only three points. These points are for tests 1, addition, 2, multiplication, and 7, identical numbers; all other tests appear to involve some form of spatial manipulation. However, each of the base planes in this example are defined by attribute vectors in the vector configuration. Thurstone's conception of simple structure provides for establishing the trait vectors and normals so that the base planes, or hyperplanes, are defined by being near trait vectors.

A number of matters related to simple structure are illustrated in the following figures. Two major classes of simple structure exist: a simple structure with a positive manifold and a structure without a positive manifold. These two classes are illustrated in Figure 7.11. A positive manifold is generated in a domain for which only zero or positive effects exist. There appear to be very few inhibitive effects in mental abilities. A major observation is that ability measures correlate zero or positively which leads to a hypothesis that there will be no negative influences of factors on performance of ability tasks. Of course, there may be a rare exception so that investigators should be alert to this possibility and allow a negative factor loading in the transformation of factors. In contrast to the ability domain is the personality measures domain. Negative relations are commonly observed between personality measures which leads to the possibilities of negative factor loadings. The factor transformation problem is much more difficult when there is not a positive manifold.

Figure 7.12 presents relations between clusters of attributes and simple structure. The sphere on the left pictures an independent cluster configuration with the attribute vectors clustered at the corners of the configuration. Each attribute is dependent on only one factor. While this is an extreme of simple structure, it is not the only possibility. A common misconception of simple structure is that Thurstone's principle is satisfied only by an independent cluster configuration. The configuration on the right pictures a simple structure with no pure attributes at the corners. The planes are well defined by points between the corners. This configuration on the right has dependent clusters, one in the middle and one part way along the bottom plane. Some proposed methods of factor analysis have used clusters to define factors, both in the extraction of factors and in the transformation of factors from an original factor matrix. The existence of clusters of attributes may have very interesting scientific meaning; however, clusters are not reliable in defining simple structures. A well defined simple structure plane will have attribute vectors in the plane stretched out from one corner to the other. Being in a plane is to be interpreted as being only a trivial distance away from the plane. The definition of a simple structure plane, or hyperplane in higher dimensions, depends not only on the number of attribute vectors in the plane but also on the diversity of these attribute vectors.

The possibility of a general factor presents a problem for simple structure. If there is a common factor for a battery of attributes, there would be no attributes having zero loadings to define the hyperplane. The upper left configuration in Figure 7.13 illustrates the case of a common factor. There are no points on the lower plane. A possibility which has been followed in a few studies is to establish the planes for all factors except for the general factor and, then, to set the plane for the general factor orthogonal to the other planes.

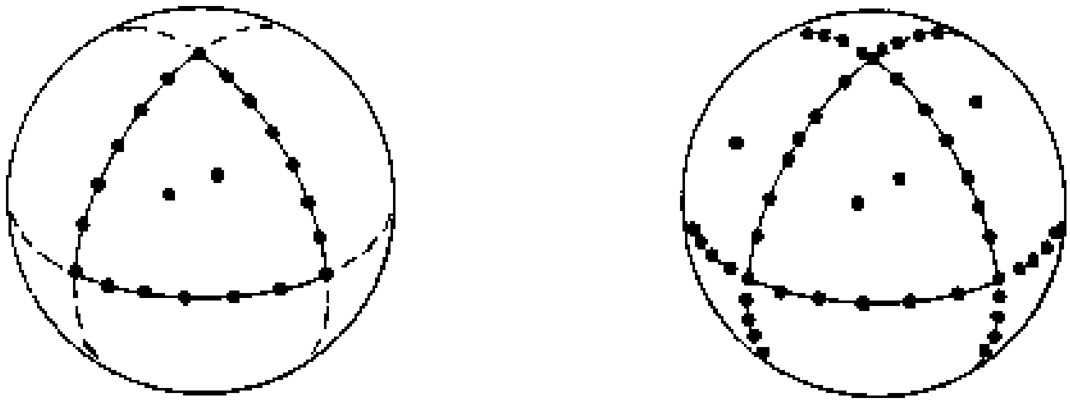

Figure 7.11: Shperical Representations of Two Classes of Simple Structure: 1. A Positive Manifold; 2. A Configuration without a Positive Manifold

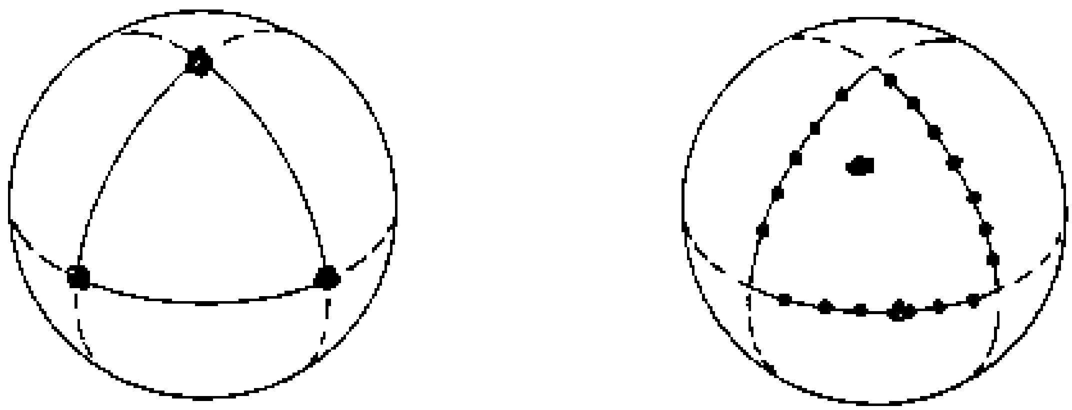

Figure 7.12: Simple Structures with Clusters of Attribute Vectors: 1. Independent Clusters;
2. Dependent Clusters

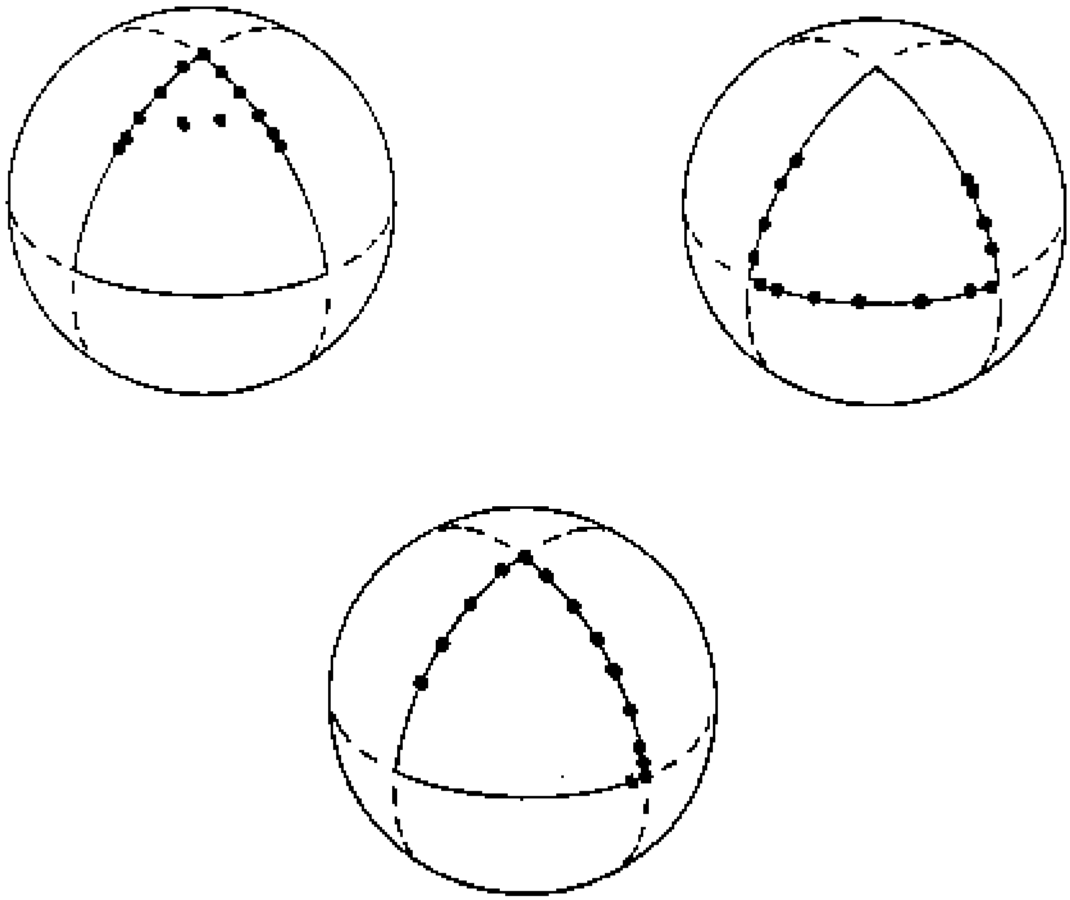

Figure 7.13: Incomplete Simple Structures: 1. A General Factor; 2. A Structure with a Corner Missing; 3. A structure with an Undefined Plane

The configuration at the upper right of Figure 7.13 presents another common type of problem: there are no points near the extreme of one factor. All attributes depend on other factors plus some dependence on the upper factor. For an example consider the case for a reasoning factor. The attributes that depend on this reasoning factor also depend upon a content type factor. There appears to be no pure reasoning tests. As a result, the hyperplanes for the other factors are less well defined by having points only part way toward the reasoning trait vector.

The bottom configuration of Figure 7.13 illustrates a difficulty derived from an incomplete battery of attributes. There are no attribute measures along the bottom plane in the direction of the left trait vector. This leaves this plane indeterminant. A solution may be available for future studies to include in the battery measures which would not depend on the factor for the upper trait but would depend upon the factor to the left side. A tentative solution for the present study would involve setting the left trait vector orthogonal to the upper trait vector.

In the preceding paragraphs we have reviewed some properties and problems of the transformation to simple structure. Many years of experience have indicated that Thursone's principal of simple structure is an extremely useful tool in the transformation of factors to meaningful solutions.