CHAPTER 10
FACTOR TRANSFORMATIONS: GRAPHICAL ROTATION

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

CHAPTER 10
FACTOR TRANSFORMATIONS: GRAPHICAL ROTATION

After a factor matrix has been extracted from a covariance matrix by one of the methods discussed in Chapters 8 and 9, the next part of exploratory factor analysis is the transformation of the factors. The basic theory for transformation of factors was discussed in Chapter 7 along with a geometric model. In this chapter procedures will be discussed for transformation of factors by graphical methods which utilize the geometric model. A simple situation involves a two factor solution. When there are more than two factors the procedures are more complex.

The procedures to be discussed represent attempts to establish transformed factors in terms of Thurstone's principle of simple structure. Such a transformed solution is not always possible for all bodies of data; there is a matter of judgment in deciding the extent to which a simple structure has been achieved. It is not a matter of a simplest structure as interpreted by some individuals but is a matter of absolute judgment whether or not a simple structure was achieved. The strength of a simple structure may be judged in terms of the number and variety of attributes that have trivial loadings on each transformed factor. As discussed in Chapter 7 there are a variety of simple structures some of which are not complete. The completeness of a structure is a function of the attributes in a battery for which an analysis is being conducted. Each of the incomplete structures poses problems in the transformation of factors and in the interpretation of the factorial results. A major goal of factor transformations is to provide bases for conjectures as to the structure of the phenomena underlying the measured attributes and the relations among these attributes.

The graphical rotations of factors operate in terms of an original factor matrix $\boldsymbol{A}$. Usually, the attributes are assumed to be scaled in terms of standard scores. In case the original factor matrix was derived from analysis of a covariance matrix, the obtained factor matrix should be scaled so as to refer to standardized scores on the attributes. Let $\widetilde{\boldsymbol{A}}$ be the original factor matrix derived from the analysis of a covariance matrix and $\boldsymbol{S}$ be the diagonal matrix of standard deviations of the attributes. Then, $\boldsymbol{A}$, the factor matrix for standard scores on the attributes is obtained by:

$$\boldsymbol{A} = \boldsymbol{S}^{-1}\widetilde{\boldsymbol{A}} \quad . \tag{10.1}$$

The discussions in this chapter will consider only factor matrices for standard scores on the attributes.

The coordinate system for transformed factors is defined by base entities, base lines for two factor studies, base planes for three factor studies, and base hyperplanes for studies having more than three factors. These base entities may be defined by normals orthogonal to the base entities. Matrix $\boldsymbol{F}$ contains as row vectors the coordinates of the normals. Since the normals are unit length vectors:

$$Diag(\boldsymbol{F}\boldsymbol{F'}) = \boldsymbol{I} \tag{10.2}$$

Once having the normals matrix $\boldsymbol{F}$, the trait matrix, with trait vectors as rows, may be obtained by a solution from equation (7.6) as:

$$\boldsymbol{T} = \boldsymbol{D}(\boldsymbol{F'})^{-1} \ . \tag{10.3}$$

By equation (7.20), the trait vectors are of unit length so that:

$$Diag(\boldsymbol{T}\boldsymbol{T'}) = \boldsymbol{I} \ . \tag{10.4}$$

Diagonal matrix $\boldsymbol{D}$ may be obtained by:

$$\boldsymbol{D} = [Diag(\boldsymbol{F}\boldsymbol{F'})^{-1}]^{-\frac{1}{2}} \ . \tag{10.5}$$

The factor correlation matrix is given by equation (7.1) as:

$$\boldsymbol{R_{bb}} = \boldsymbol{T}\boldsymbol{T'} \ . \tag{7.1}$$

An alternative formula for the factor correlation matrix is:

$$\boldsymbol{R_{bb}} = \boldsymbol{D}(\boldsymbol{F}\boldsymbol{F'})^{-1}\boldsymbol{D} \ . \tag{10.6}$$

There are four types of coefficients for the attribute vectors in the transformed factor structure. Only three of these will concern us here: matrix $\boldsymbol{G}$ contains the orthogonal projections of the attribute vectors on the normals; matrix $\boldsymbol{B}$ of factor weights of the modeled attributes; matrix $\boldsymbol{Q}$ of covariances of modeled attributes with the traits. Matrix $\boldsymbol{G}$ is given by equation (7.9)

$$\boldsymbol{G} = \boldsymbol{A}\boldsymbol{F'} \ . \tag{7.9}$$

Matrix $\boldsymbol{B}$ is given by equation (7.4)

$$\boldsymbol{B} = \boldsymbol{A}\boldsymbol{T}^{-1} \ . \tag{7.4}$$

Matrix $\boldsymbol{Q}$ is given by equation (7.5)

$$\boldsymbol{Q} = \boldsymbol{A}\boldsymbol{T'} \ . \tag{7.5}$$

Graphical transformation of factors uses the projections of the attribute vectors on the normals as measures of importance of the transformed factors to the modeled attributes. Note that these coefficients are proportional, factor by factor, to the factor weights in matrix $\boldsymbol{B}$. By equation (7.10):

$$\boldsymbol{G} = \boldsymbol{B}\boldsymbol{D} \ . \tag{7.10}$$

An important point is that each column of $\boldsymbol{G}$ depends only upon the corresponding column of $\boldsymbol{F}$ while each column of factor weights in $\boldsymbol{B}$ depends upon the location of all the normals in matrix $\boldsymbol{F}$. This independence permits the shifting of each normal independently from the location of the other normals.

10.1. Graphical Rotation for Two Factor Studies

For a two factor study a single graph provides a view of the configuration of attribute vectors and transformed coordinate vectors. This single graph provides a convenient view of the entire transformed factorial structure. For a three factor study a unit sphere may be used to provide a view of the configuration of attribute vectors and transformed coordinate vectors as illustrated in Chapter 7, Figure 7.8; however, this is quite inconvenient and is not used in practice. The two dimensional representation in a graph provides a convenient opportunity to view the configuration of attribute vectors and the transformed coordinate vectors. The combination of attribute vectors and transformed coordinate vectors is termed a structure. Thurstone's simple structure is a structure well defined by many attribute vectors near the base entities, base lines for two dimensional studies.

Graphical rotation of factors for a two factor study will be illustrated by an analysis of a selected battery of verbal comprehension and word fluency tests taken from the Thurstone & Thurstone (1941) factorial studies of intelligence. The correlation matrix for the selected nine tests is given in Table 10.1 and the principal factors matrix is given in Table 10.2. This principal factors matrix is the given matrix $\boldsymbol{A}$ for which the factors are to be transformed. Figure 10.1 provides the graph of the attribute vector configuration. Each row of matrix $\boldsymbol{A}$ contains the coordinates of the terminus of an attribute vector on the principal factor axes. In the graph of Figure 10.1, the horizontal axis represents coordinates on principal factor 1 while the vertical axis represents coordinates on principal axis 2. There is a point on this graph for the terminus of each attribute vector. For clarity of the graph, the vectors are not drawn. Each vector emanates from the origin to its terminus represented by a point. In this illustration the points form two clusters. Such a configuration is not required for a simple structure. It would be acceptable to have a few points in the middle between two clusters. Also, the points do not have to form a cluster; rather, a rough radial line of points could replace a cluster.

In Figure 10.2 the lines of the transformed factor system have been added to the configuration of attribute points shown in Figure 10.1. This system can be thought of as defined by base lines, there being one such base line for each transformed factor. Consider Base Line 1 and Base Line 2 drawn on Figure 10.2. Each base line is restricted to passing through the origin.

Table 10.1

Correlation Matrix for V & W Tests Example

Variable	1	2	3	4	5	6	7	8	9
1	1.000								
2	.769	1.000							
3	.718	.681	1.000						
4	.730	.661	.672	1.000					
5	.227	.189	.280	.241	1.000				
6	.296	.219	.311	.311	.554	1.000			
7	.237	.212	.313	.245	.461	.479	1.000		
8	.243	.226	.348	.290	.506	.530	.520	1.000	
9	.304	.291	.374	.306	.408	.425	.514	.473	1.000

*Selected from 66 test study by Thurstone and Thurstone (1941). Factorial studies of intelligence; sample size = 710.

Table 10.2

Principal Factors Matrix for V & W Tests Example

Variable		1	2
Chicago Reading Tests, Vocabulary	1	.765	-.435
Chicago Reading Tests, Sentences	2	.707	-.437
Completion	3	.766	-.281
Same or Opposite	4	.727	-.337
Prefixes	5	.530	.425
Suffixes	6	.583	.401
First and Last Letters	7	.552	.415
First Letters	8	.583	.420
Four-Letter Words	9	.575	.294

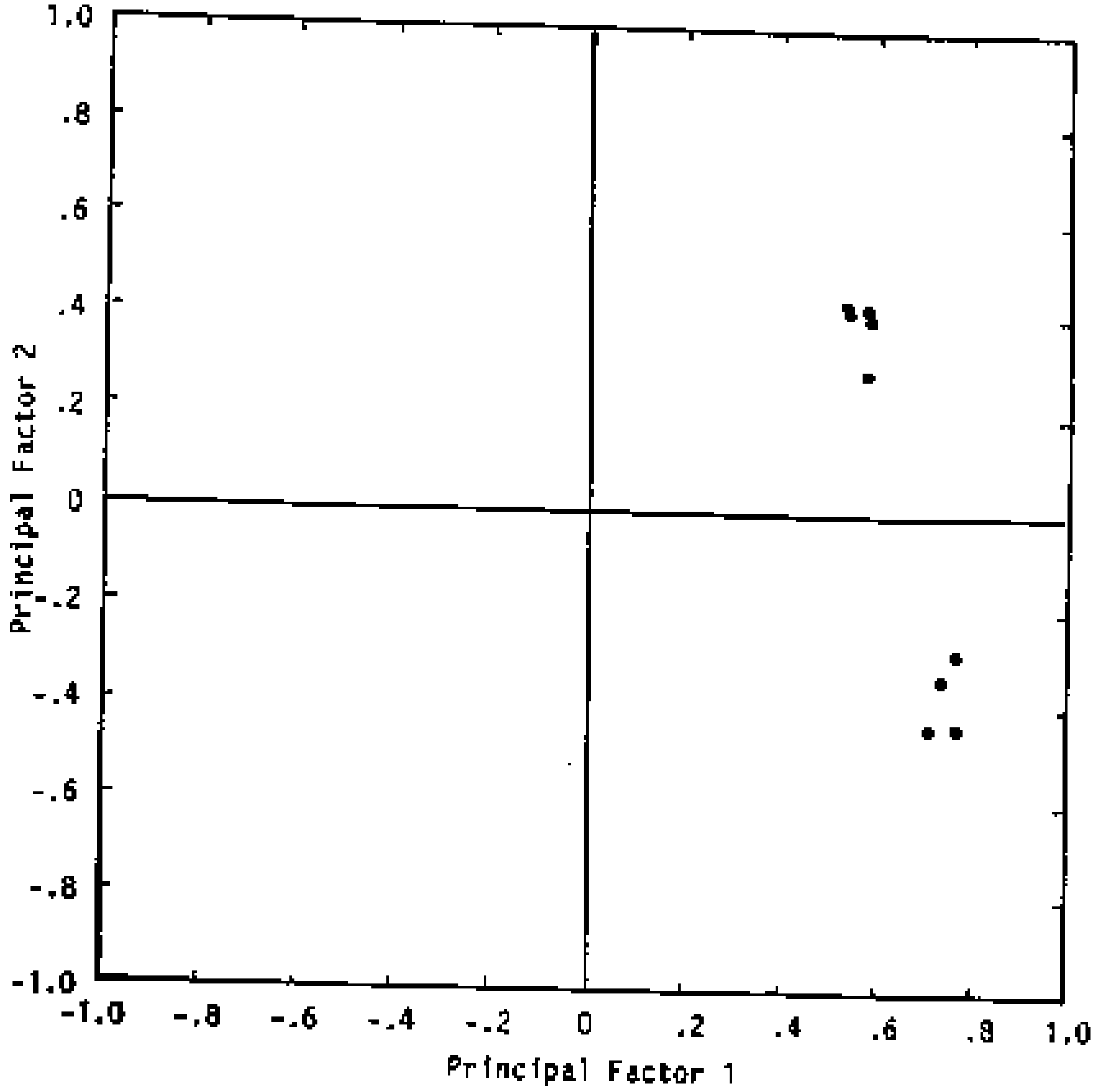


Figure 10.1 Principal Factors Configuration for V & W Tests Example.

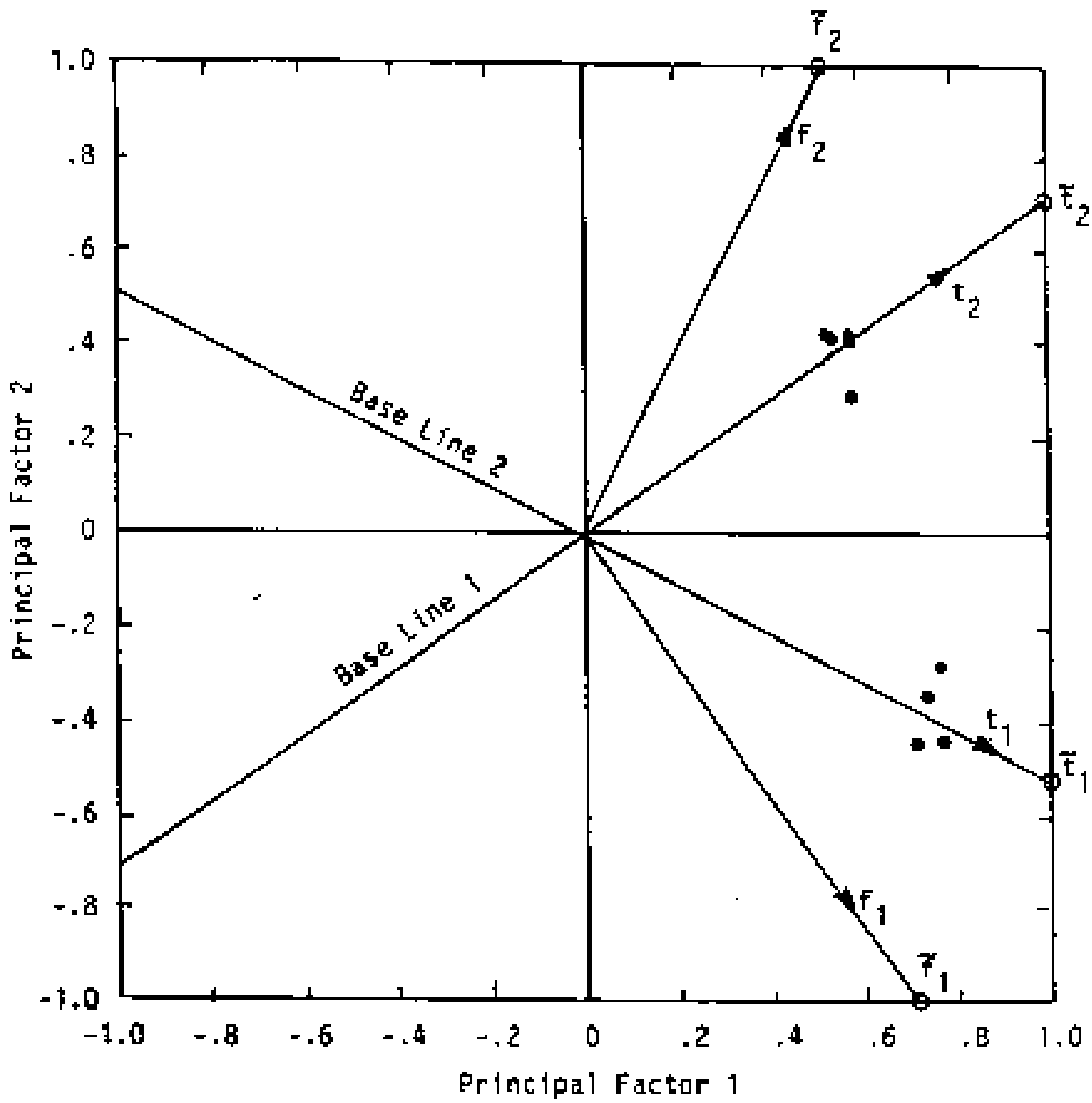


Figure 10.2 Factor Structure for V & W Tests Example.

In graphical rotation of factors subjective judgments are involved in drawing the base lines. For a simple structure each base line should pass near to a number of attribute points which may be in a rough radial line, there being no restriction that the points be in a cluster as in the illustration. Each base line is important in that it forms the zero for measuring influence of the transformed factor on each attribute. For each attribute having a point at a trivial distance from the base line the factor has a trivial influence. As the distances to points increase, influences of the factor on the attributes increase. While theoretic statements emphasize points at zero distances from a base line, in practice points are considered as being "in a base line" when these points are within a small interval from the base line. Frequently, this interval is considered to have a width of $\pm.10$. Subjective judgment is involved in seeing a radial line of points in the configuration which lie within such an interval from a possible base line and in drawing the base line.

Base Line 1 is drawn on Figure 10.2 from the lower left through the origin and the cluster of points in the upper right quadrant to a point $\tilde{t}_2$ which was chosen to have coordinates easily read from the graph. It is not necessary that such points be at a unit distance from the origin; when necessary, unit length vectors can be computed. The coordinates of point $\tilde{t}_2$ are (1.00, .71) which are recorded in Table 10.3 as line 2 of the Raw Traits Matrix $\tilde{\boldsymbol{T}}$. The normal to Base Line 1 is orthogonal to this base line and emanates from the origin to point $\tilde{f}_1$ at the lower right. This normal could have been directed toward the upper left but was chosen to be directed to the lower right so that the attribute points distant from the base line would be on the positive side of the base line. Coordinates of $\tilde{f}_1$ are (.71, -1.00) which are obtained from the coordinates of point $\tilde{t}_2$ by interchanging the coordinates of this point and changing the sign of one of the coordinates. Choice of which coordinate for which the sign is to be changed is according to the direction the normal is to take. This interchange of coordinates and change in the sign of one of the coordinates makes the scalar product equal to zero of vectors emanating from the origin to points $\tilde{t}_2$ and $\tilde{f}_1$. The coordinates of point $\tilde{f}_1$ are recorded in Table 10.3 in line 1 of the Raw Normals Matrix $\tilde{\boldsymbol{F}}$.

Base Line 2 is drawn on Figure 10.2 from the upper left through the origin and the cluster of attribute points at the lower right to point $\tilde{t}_1$ with coordinates (1.00, -.51) which are recorded in line 1 of the Raw Traits Matrix $\tilde{\boldsymbol{T}}$ in Table 10.3. The normal to Base Line 2 is orthogonal to this base line, emanates from the origin to point $\tilde{f}_2$ with coordinates (.51, 1.00) which are recorded in line 2 of the Raw Normals Matrix $\tilde{\boldsymbol{F}}$ in Table 10.3. The coordinates for $\tilde{f}_2$ were obtained by interchanging the coordinates of point $\tilde{t}_1$ and changing the sign of one of these coordinates. Choice of which coordinate to have its sign changed is made so that the points distant from the base line will be on the positive side of the base line.

Once the raw vectors have been obtained in the Raw Traits Matrix $\tilde{\boldsymbol{T}}$ and the Raw Normals Matrix $\tilde{\boldsymbol{F}}$ the unit length vectors are to be obtained in Traits Matrix $\boldsymbol{T}$ and Normals

Table 10.3

Graphical Transformation Matrices for V & W Tests Example

Raw Traits Matrix $\tilde{T}$

	1	2	L^2	L
1	1.00	-.51	1.2601	1.1225
2	1.00	.71	1.5401	1.2264

Traits Matrix T

	1	2
1	.891	-.454
2	.815	.579

Raw Traits Matrix $\tilde{F}$

	1	2	L^2	L
1	.71	-1.00	1.5041	1.2264
2	.51	1.00	1.2601	1.1225

Normals Matrix F

	1	2
1	.579	-.815
2	.454	.891

Cosines of Angles Between Normals FF'

	1	2
1	1.000	-.463
2	-.463	1.000

Factor Correlation Matrix R_{bb} = TT'

	1	2
1	1.000	.463
2	.463	1.000

Cosines of Angles Between Normals and Trait Vectors FT'

	1	2
1	.886	.000
2	.000	.886

Matrix $\boldsymbol{F}$. Let $\widetilde{\mathbf{x}}_{\boldsymbol{k}}$ with entries $\widetilde{\mathrm{x}}_{jk}$ represent each of the raw vectors. The normalization process can be symbolized by computing the squared length, $\widetilde{\mathrm{L}}_{k}^{2}$, of the vector by:

$$\widetilde{\mathrm{L}}_{k}^{2} = \sum_{j=1}^{r} \widetilde{x}_{jk}^{2} \quad . \tag{10.7}$$

Each element in the vector is divided by the length of the vector to produce the unit length, or normalized vector:

$$\boldsymbol{x}_{\boldsymbol{jk}} = \widetilde{\boldsymbol{x}}_{\boldsymbol{jk}} / \widetilde{\mathrm{L}}_{k} \quad . \tag{10.8}$$

This procedure is applied to each of the raw vectors in matrices $\widetilde{\boldsymbol{T}}$ and $\widetilde{\boldsymbol{F}}$ to yield the Traits Matrix $\boldsymbol{T}$ and $\boldsymbol{F}$. Normals Matrix $\boldsymbol{F}$ in Table 10.3. These unit length vectors are shown by arrows on Figure 10.2.

The remaining matrices in Table 10.3 contain scalar products among the vectors in matrices $\boldsymbol{T}$ and $\boldsymbol{F}$. Since the vectors in matrices $\boldsymbol{T}$ and $\boldsymbol{F}$ have been normalized to unit length vectors, these scalar products between these vectors are cosines of the angles between these vectors. The first of these scalar product matrices is the product $\boldsymbol{FF'}$ and contains the cosines of the angles between the normals. The diagonal entries of unity are the lengths of the normal vectors and must be unity since these vectors were normalized. The next matrix is the scalar product matrix among the trait vectors and is obtained by the product $\boldsymbol{TT'}$. By equation (7.1) this is the factor correlation matrix $\boldsymbol{R_{bb}}$. In the two dimensional factor rotation case the off-diagonal entries in $\boldsymbol{R_{bb}}$ must be equal to the negative of the off-diagonal entries in $\boldsymbol{FF'}$. The bottom matrix contains the scalar products of the vectors in $\boldsymbol{F}$ and $\boldsymbol{T}$. This is a diagonal matrix due to the construction of the normals vectors being orthogonal to the base planes which contain the trait vectors. This is the diagonal matrix $\boldsymbol{D}$ as used in equations (7.6) and (10.3).

Table 10.4 gives the matrices for coefficients of the modeled attributes on the transformed factors. Matrix $\boldsymbol{G}$ contains the projections of the modeled attributes on the normals. These are the perpendicular distances of the attribute points from the base lines and are computed by equation (7.9); they are the scalar products of the attribute vectors with the normals. A convenient interpretation for the projection of an attribute vector on a normal is that it represents the contribution of a transformed factor to the attribute which is independent of the other transformed factors. It depends only on the location of the particular base line and not on the location of the other base line. In matrix $\boldsymbol{G}$ for the illustration the first four attributes have high coefficients on the first transformed factor. These are attributes for which their points were some distance from Base Line 1 which was drawn to be near the points for the last five attributes. Base Line 1 was drawn to be near these last five points. Note that the projections of these last five points are in the interval of $\pm.10$. Factor 1 is a good representation of a simple structure factor.

Table 10.4

Graphical Transformed Factor Matrices for V & W Test Example

Matrix G of Projections on Normals

	1	2
1	.798	-.040
2	.766	-.068
3	.673	.098
4	.696	.030
5	-.040	.619
6	.011	.622
7	-.019	.620
8	-.005	.639
9	.093	.523

Factor Weight Matrix B

	1	2
1	.900	-.045
2	.864	-.077
3	.759	.110
4	.785	.034
5	-.045	.699
6	.012	.702
7	-.021	.700
8	-.006	.721
9	.105	.590

Matrix Q of Covariances of Modeled Attributes with Traits

	1	2
1	.879	.372
2	.828	.323
3	.810	.482
4	.801	.398
5	.279	.678
6	.337	.708
7	.303	.690
8	.329	.719
9	.379	.639

For factor 2, the first four attributes have projections in the $\pm.10$ interval which provides a nice representation of a simple structure factor. The locations of the base lines is dependent on the points having trivial projections on the normals.

Factor Weight Matrix $\boldsymbol{B}$ is the second matrix of Table 10.4 and was computed by equation (7.4). Its relation to the matrix of projections is given by equation (7.10) which can be rewritten as:

$$\boldsymbol{B} = \boldsymbol{G}\boldsymbol{D}^{-1} \ . \tag{10.9}$$

The factor weights are proportional to the projections on the normals, the constants of proportionality being the reciprocals of the diagonal entries in matrix $\boldsymbol{D}$. For uncorrelated transformed factors matrix $\boldsymbol{D}$ is an identity matrix and matrices $\boldsymbol{G}$ and $\boldsymbol{B}$ are equal. As the factors become more correlated, the entries in $\boldsymbol{D}$ become smaller so that the entries in $\boldsymbol{B}$ become larger for the same sized entries in $\boldsymbol{G}$. As with multiple regression, when the independent variables are highly correlated the regression weights are large and do not represent contributions to the dependent variable. This is true for the factor weights.

Matrix $\boldsymbol{Q}$ of Table 10.4 contains the covariances of Modeled Attributes with the traits. These are scalar products of the attribute vectors in matrix $\boldsymbol{A}$ with the trait vectors as given in equation (7.5). Equations (7.1), (7.4), and (7.5) may be combined to produce:

$$\boldsymbol{Q} = \boldsymbol{B}\boldsymbol{R}_{bb} \ . \tag{10.10}$$

It is apparent from this equation that the coefficients in $\boldsymbol{Q}$ combine both the factor weights and the factor correlations. Therefore, these coefficients do not represent the independent contributions of the factors to the modeled attributes. Note that: for uncorrelated factors matrix $\boldsymbol{D}$ is an identity matrix so that matrices $\boldsymbol{G}$, $\boldsymbol{B}$, and $\boldsymbol{Q}$ are equal. Some investigators prefer this type transformation; however, when base lines are well defined by radial lines of points, the correlated factors transformation would provide a preferred simple structure solution.

10.2. Graphical transformation of Factors for Three and More Factor Studies.

For a three factor study, a vector model of the configuration of points can be constructed and a transformed coordinate system of planes, normals and trait vectors inserted. As an alternate, Thurstone (1947) described a procedure to plot the termini of attribute vectors extended to unit length on the surface of blackboard sphere and to draw in the coordinate planes as great circles. The coordinates of the trait vectors and normals could be read on this sphere. Such a procedure was quite unwieldy and not used except as a demonstration. Further, such a procedure was not available for studies involving four or more factors. A usable procedure was to make a set of two dimensional graphs plotting projections on normals for all pairs of factors. This procedure involved a series of rotations with each rotation made from a set of graphs for tentative transformed normals. Projections on a new tentative normals were computed and a new set of graphs were made. This procedure was continued until a satisfactory set of graphs were obtained. Details for this procedure are described in the following paragraphs with an illustration using the nine mental tests example.

Graphical transformations start from some initial transformation. In the case of the illustration the procedure was started from the principal factors for the nine mental tests example. This implied an initial normals matrix equal to an identity matrix. Table 10.5 gives the initial normals matrix and projections on these normals (which equal the loadings on the principal axes). Frequently some one of the analytical transformation methods (to be described in Chapter 11) is used to provide the initial normals matrix and projections on the normals. A normal VARIMAX transformation provides a good initial normals matrix and projections on the normals. Sometimes the computer output does not give the normals matrix, only the factor weights, matrix $\boldsymbol{B}$, and/or the projections on the normals, matrix $\boldsymbol{G}$, being given along with the correlations among the factors, matrix $\boldsymbol{R_{bb}}$. In this case the initial normals matrix, $\boldsymbol{F}_0$, must be computed by the appropriate one of the following formulas.

$$\boldsymbol{F}_0 = \boldsymbol{G'A}(\boldsymbol{A'A})^{-1} \quad . \tag{10.11}$$

$$\boldsymbol{F}_0 = \boldsymbol{DB'A}(\boldsymbol{A'A})^{-1} \tag{10.12}$$

where

$$\boldsymbol{D} = [Diag(\boldsymbol{R_{bb}^{-1}})]^{-\frac{1}{2}} \quad . \tag{7.8}$$

Matrix $\boldsymbol{G}_0$ of projections on the normals is given by:

$$\boldsymbol{G}_0 = \boldsymbol{AF}_0' \quad . \tag{7.9}$$

An initial set of factor plots is made using projections on the normals in matrix $\boldsymbol{G}_0$. Each pair of factors is used in making one of the plots in the set; thus, there are r(r $-$ 1)/2 plots in the set. For the three factor, nine mental tests example there are three plots in Figures 10.3a, 10.3b, 10.3c. The attribute points are labeled with the attribute numbers for convenience in cross referencing between the several plots in the set. A base line is drawn using subjective judgment for each of the transformed factors on one or another of the plots in the set. Each base line should pass near a number of the attribute points. With aptitude and intelligence measures each base line should be near one side of the configuration of points in accordance with an assumption that all factor weights in these cases should be positive. This is a very powerful principal termed a positive manifold.

Table 10.5

Initial Graphical Transformation
Nine mental Tests Example

Normals Matrix F_0

	1	2	3
1	1.000	.000	.000
2	.000	1.000	.000
3	.000	.000	1.000

Projections on Normals

Matrix G^*_0

Variable		1	2	3
Addition	1	.42	.36	.28
Multiplication	2	.47	.54	.16
Three-Higher	3	.61	.16	.19
Figures	4	.54	-.46	.09
Cards	5	.63	-.48	.05
Flags	6	.59	-.41	.06
Identical Numbers	7	.48	.40	-.21
Faces	8	.61	.00	-.29
Mirror Reading	9	.59	.16	-.25

*Principal factor of R with SMC's in diagonal.

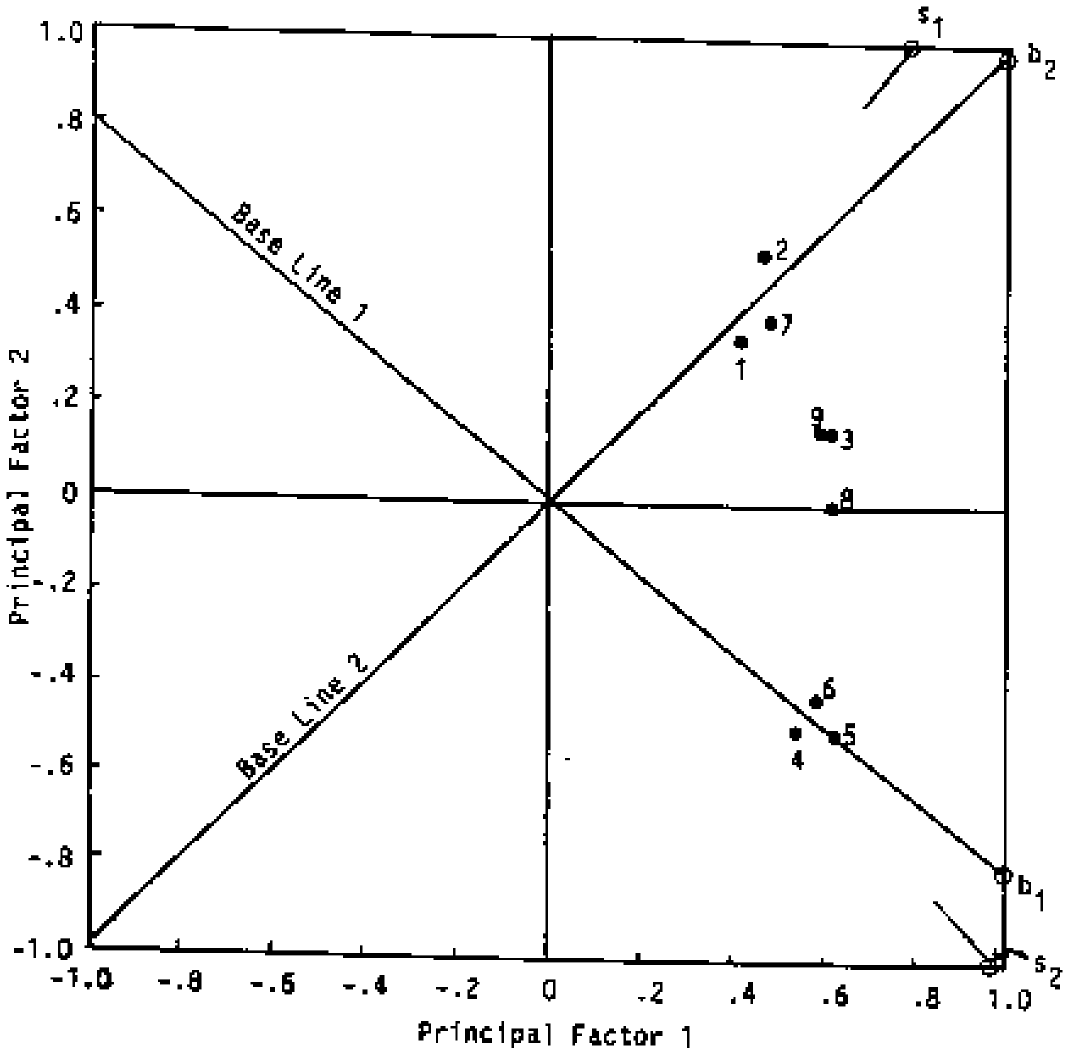


Figure 10.3a Nine Mental Tests: Factor Plot for Principal Factors 1 and 2.

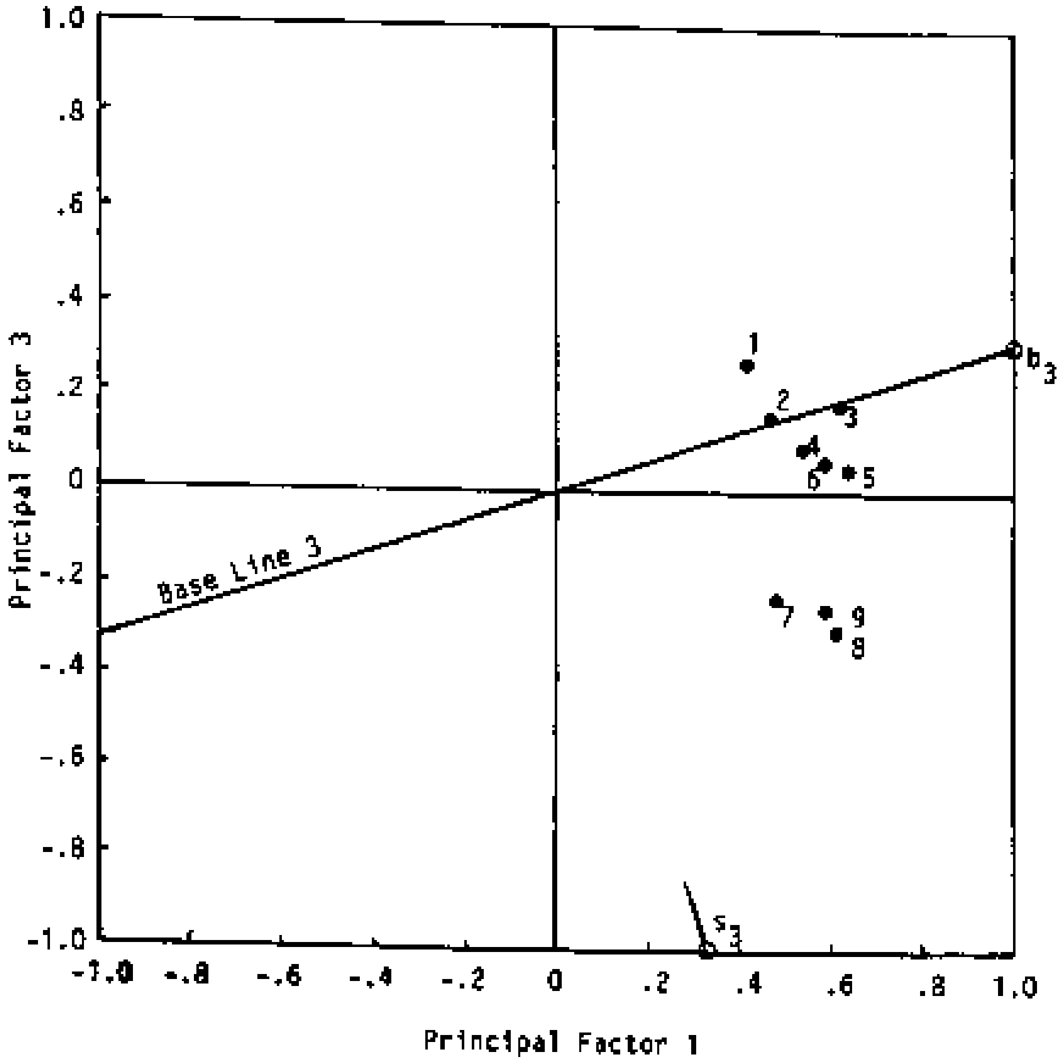


Figure 10.3b Nine Mental Tests: Factor Plot for Principal Factors 1 and 3.

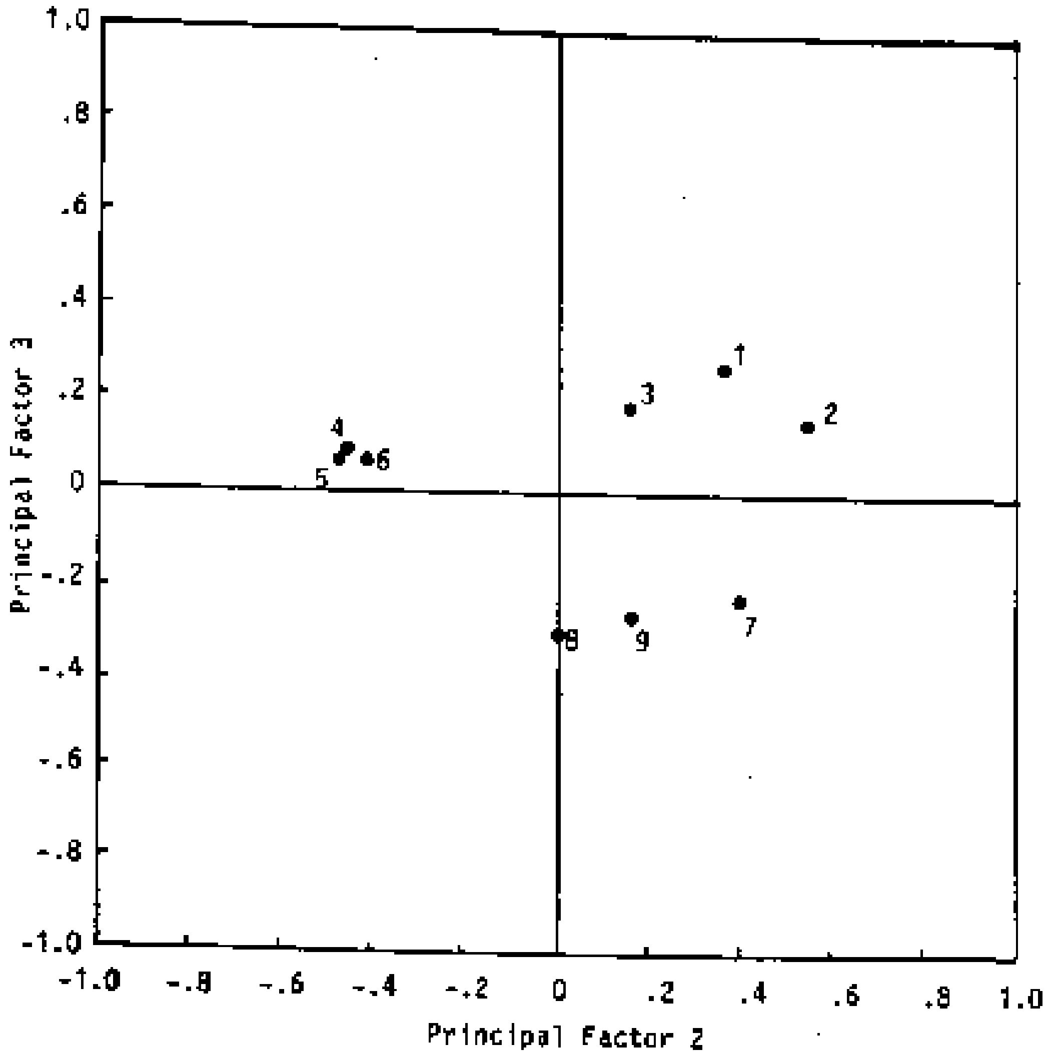


Figure 10.3c Nine Mental Tests: Factor Plot for Principal Factors 2 and 3.

Base Line 1 for transformed factor 1 is drawn on the plot for factors 1 and 2 on Figure 10.3a to be near points 4, 5, 6. Point b_1 is chosen on Base Line 1 to have conveniently read coordinates of (1.00, -.80) and point s_1 is indicated to be on a line from the origin orthogonal to Base Line 1. The coordinates of s_1 , (.80, 1.00), have the coordinates of b_1 reversed with one of them reversed in sign. Choice as to which coordinate to reverse in sign is dictated by the location desired for point s_1. The desired point may be plotted on the graph as shown on Figure 10.3a. This direction of the point was chosen so that the majority of the attribute points would be on the positive side of Base Line 1.

Base Line 2 was drawn on the plot between factors 1 and 2 on Figure 10.3a from the lower left through the origin and near points 1, 2, 7 to point b_2 with coordinates (1.00, .98). Point s_2 is at the lower right on a line from the origin orthogonal to Base Line 2. The coordinates of s_2 are (.98, -1.00) which are the coordinates of b_2 interchanged and the sign of one coordinate reversed in sign. The direction to the lower right was chosen such that the attribute points not near Base Line 2 would be on the positive side of Base Line 2. Note that the cluster of points 3, 8, 9 was ignored in drawing base lines since this cluster is in the middle of the configuration so that some attribute points would be on the negative side of such a base line while other attribute points would be on the positive side. To choose such a base line would violate the principal of a positive manifold.

Base Line 3 was drawn on the plot for factors 1 and 3 on Figure 10.3b near a loose cluster of all attribute points except for points 7, 8, 9. Point b_3 has coordinates (1.00, .32). Point s_3 has coordinates (.32, -1.00).

The raw trait vectors, $\widetilde{\mathbf{t}_k}$, of the plot for a two factor study illustrated on Figure 10.2 have been replaced by base line marked points b_k . Likewise, the raw normals vectors, $\widetilde{f_k}$, have been replaced by shift vectors $\mathbf{s}_k$ which operate to shift the normals vectors to new raw normals vectors. These shift vectors are recorded as rows in a shift matrix $\boldsymbol{S}$. Consider Table 10.6; the three shift vectors are defined by points s_1 , s_2 , s_3 from the factor plots in Figures 10.3a, 10.3b. The coordinates of these shift vectors are recorded in Shift Matrix $\boldsymbol{S}_1$. The new raw normals matrix, $\widetilde{F_1}$, is obtained by:

$$\widetilde{F_1} = \boldsymbol{S}_1\boldsymbol{F}_0 \tag{10.13}$$

The vectors in raw normals matrix $\widetilde{\boldsymbol{F}_1}$ are normalized to unit length vectors in the normals matrix $\boldsymbol{F}_1$. Matrix product $\boldsymbol{F}_1\boldsymbol{F}_1'$ containing the cosines of the angles between the normals appears at the lower left of Table 10.6. These cosines of angles will be important to consider in the next set of graphs. The projections on the normals, matrix $\boldsymbol{G}_1$, is computed by:

$$\boldsymbol{G}_1 = \boldsymbol{A}\boldsymbol{F}_1' \ . \tag{7.9}$$

Table 10.6

Graphical Transformation 1
Nine Mental tests Example

Shift Matrix S_1

	1	2	3
1	.80	1.00	.00
2	.98	-1.00	.00
3	.32	.00	-1.00

Raw Normals Matrix $\tilde{F}_1$

	1	2	3	L^2	L
1	.800	1.000	.000	1.6400	1.2806
2	.980	-1.000	.000	1.9604	1.4001
3	.320	.000	-1.000	1.1024	1.0500

Normal Matrix F_1

	1	2	3
1	.625	.781	.000
2	.700	-.714	.000
3	.305	.000	-.952

Cosines of Angles between
Normals F_1F_1'

	1	2	3
1	1.00	-.12	.19
2	-.12	1.00	.21
3	.19	.21	1.00

Projections on Normals
Matrix G_1

	1	2	3
1	.54	.04	-.14
2	.72	-.06	-.01
3	.51	.31	.01
4	-.02	.71	.08
5	.02	.78	.14
6	.05	.71	.12
7	.61	.05	.35
8	.38	.43	.46
9	.49	.30	.42

After the initial graphical transformation is accomplished new factor plots are made using the projections on the normals in matrix $\boldsymbol{G}_1$. The normals are no longer orthogonal as shown in matrix $\boldsymbol{F}_1\boldsymbol{F}_1'$ which contains the cosines of the angles between the normals. A true representation of the configuration would require plotting on oblique axes. Consider Figure 10.4 for an illustration of the procedure to make these plots. Two normals, $\mathbf{f}_1$ and $\mathbf{f}_2$, are shown at an angle of 120° (cosine = -.50) from each other. Zero Line 1 is drawn at a right angle to normal vector $\mathbf{f}_1$. This is the locus of all points which have a zero projection on normal $\mathbf{f}_1$. Similarly, Zero Line 2 is drawn at a right angle to the vector for normal $\mathbf{f}_2$. Point j has projections on these normals of .3 and .5 . The projection of .3 of j on normal $\mathbf{f}_1$ is measured off on this normal and a projection line is drawn at a right angle to the normal. This projection line is parallel to Zero Line 1 and is the locus of all points which have a projection of .3 on normal $\mathbf{f}_1$. Similarly, the projection of .50 of point j on normal $\mathbf{f}_2$ is measured off on this normal and a projection line is drawn at a right angle to this normal. Point j is located at the intersection of the two projection lines. However, this procedure is quite tedious and need not be resorted to except when two normals are very oblique. The factor plots may be made on orthogonal axes with very little distortion except when two normals are very oblique (cosine of the angle between them more discrepant from zero than some value of $\pm .70$). Figure 10.5 provides an illustration of the distortion when the cosine of the angle between the normals is -.50. Eight points are on the sides of a square when an oblique plot is made as in the upper plot of Figure 10.5. When these points are plotted using orthogonal axes as in the lower figure the square is skewed to a diamond shape. Distances from upper left to lower right are increased while distances from lower left to upper right are decreased. However, all straight lines remain straight lines. Zero Line 1 has been rotated to coincide with normal vector $\mathbf{f}_2$; also, Zero Line 2 has been rotated to coincide with normal vector $\mathbf{f}_1$. Distances of points from these zero lines have remained unchanged in the transformation from oblique axes to orthogonal axes. While normal vectors appear to move around with reference to the configuration of points, the zero lines appear to have a constant relation to the configuration of points. Consequently, there is little effect on the distances of points from new base lines when these base lines are near zero lines. Experience has shown that the distortion from using orthogonal axes has little effect on subjective judgments in driving new base lines near lines of points.

The factor plots for rotation 1 of the nine mental tests example were made from the projections on the normals in matrix $\boldsymbol{G}_1$ of Table 10.6. These plots were studied for improvements in points near each of the zero lines. These improvements result in shifts of the normals. In the plot between factors 1 and 3 there are no points distance from the origin near the zero line for factor 1. This lead to a Base Line 1 being drawn among points 7, 8, 9. Then coordinates of base point b_1 were read from the graph and the coordinates of shift point s_1

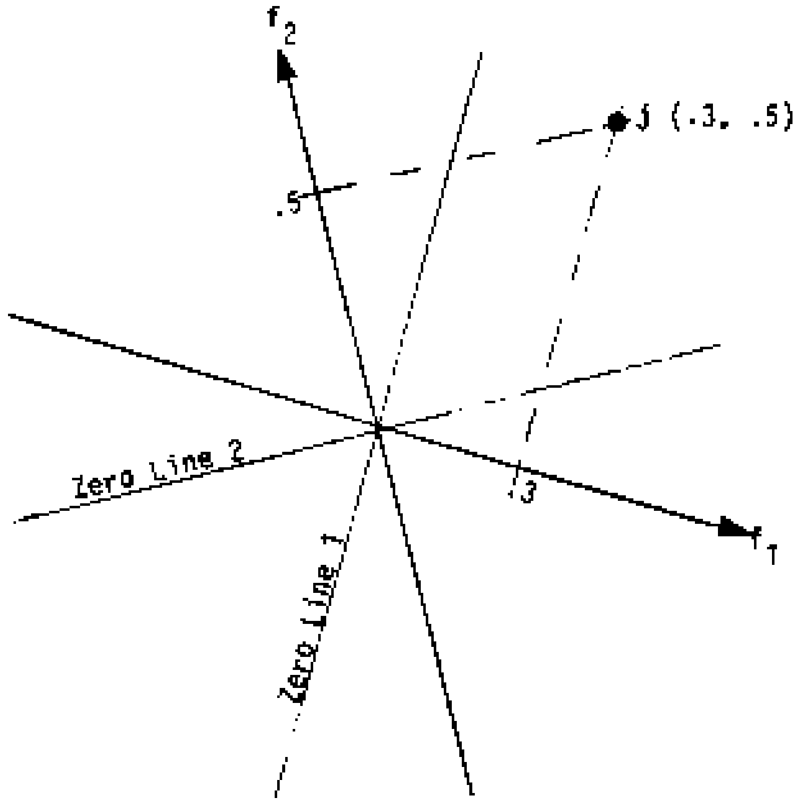
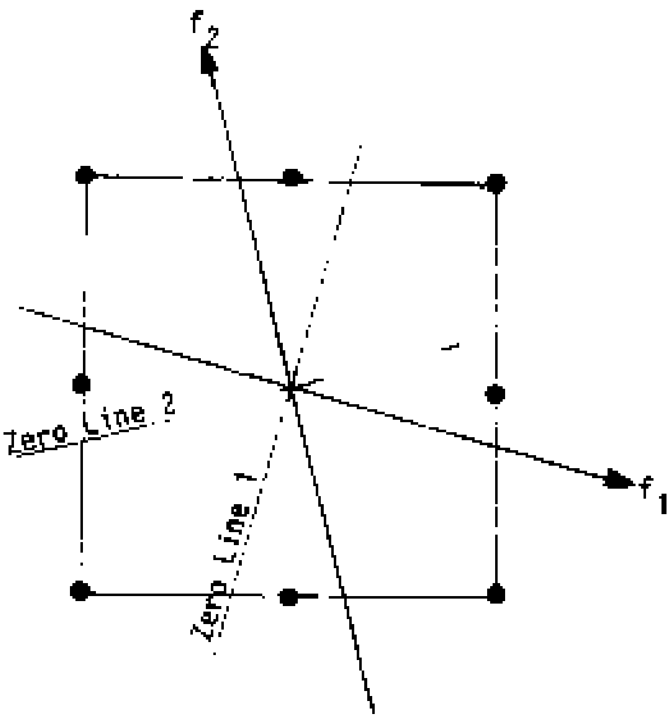


Figure 10.4 Illustration of plotting a point on oblique axes; angle between axes = 120°, cosine = -.5.



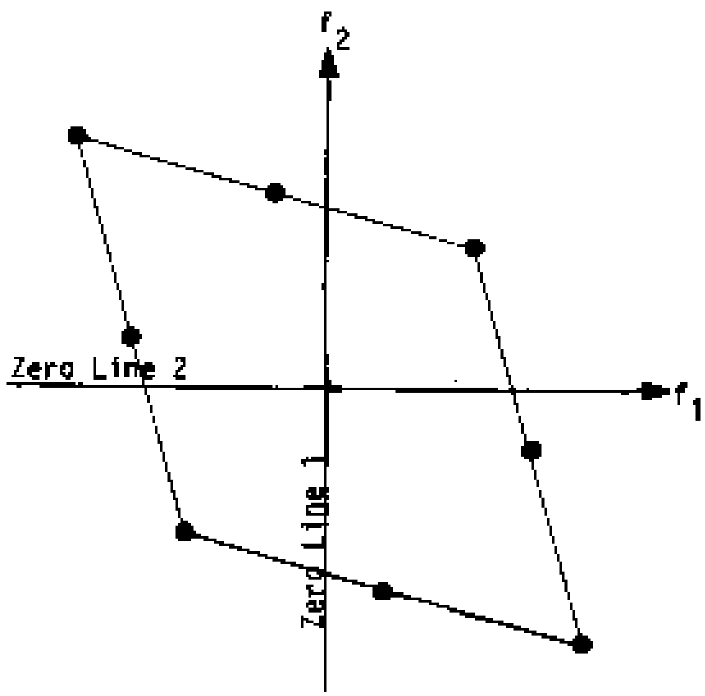


Figure 10.5 Distortion of a configuration plotted on orthogonal axes when the normals are oblique.

obtained as previously described. The coordinates of s_1 are recorded in the first row of matrix $\boldsymbol{S_2}$ in Table 10.7. A more problematic situation occurs on Figure l0.6c for the zero line for factor 2. Point 7 is near the zero line for factor 2. However, points 7, 8, 9 might form a rather wide cluster. A decision was made to try having such a cluster centered at zero for factor 2. Base Line 2 was drawn through the approximate center of this cluster. For factor 3 on the plot for factors 2 and 3 of Figure 10.6c there is a very distinct improvement resulting drawing Base Line 3 through the cluster of points 4, 5, 6 . Shift points s_2 and s_3 were developed from base points b_2 and b_3. The coordinates of s_2 and s_3 are recorded in rows of shift matrix $\boldsymbol{S_2}$ in Table 10.7.

The raw normals matrix, $\widetilde{\boldsymbol{F_2}}$ for rotation 2 is given in Table 10.7 and was computed by:

$$\widetilde{\boldsymbol{F_2}} = \boldsymbol{S_2 F_1} \ .$$

Rows of $\widetilde{\boldsymbol{F_2}}$ were normalized to unit length in normals matrix $\boldsymbol{F_2}$ and the matrix, $\boldsymbol{F_2 F_2'}$, was computed. The matrix of projections on the revised normals, $\boldsymbol{G_2}$ was computed by:

$$\boldsymbol{G_2} = \boldsymbol{A F_2'} \ .$$

Factor plots for rotation 2 are given in Figures 10.7a, 10.7b, 10.7c and were inspected for further adjustments of the zero lines to be near radial lines of points. A small adjustment for factor 1 appeared desirable on the plot for factors 1 and 2 in Figure 10.7a. Also, a small adjustment for factor 3 appeared desirable on the plot for factors 1 and 3 in Figure 10.7b. A larger adjustment for factor 2 was decided upon, see Figure 10.7c for the plot between factors 2 and 3. This adjustment approximately reverses a decision made for rotation 1 on Figure 10.6c. At that time it was noted that points 7, 8, 9 might form a loose cluster. The further view given in Figure 10.7c lead to a judgment that these points were too widely separated to form such a cluster. Two other matters were considered in deciding to pass the new Base Line 2 through point 7: first was the principle of a positive manifold; second was a matter of interpretation. With the decision to pass Base Line 2 through point 7, all points except point 2 were now positive with point 2 only trivially negative. This satisfies the positive manifold principle. Use of interpretation to guide factor transformations greatly weakens the strength of conclusions derived from a study. However, in the case of the example, reference to the attributes named in Table 10.5 indicates that factor 2 could be some sort of spatial factor with attributes 8 and 9. Faces and Mirror Reading, having small positive loadings. As noted above, this use of interpretations to guide factor transformations weakens conclusions which may be made from factor studies. Only conjectures as to the nature of factors appear to be justified. Follow up studies involving new measures based upon these tentative conclusions are imperative to strengthen interpretations. A program of studies should lead to confirmatory studies based on interpretations developed in the program.

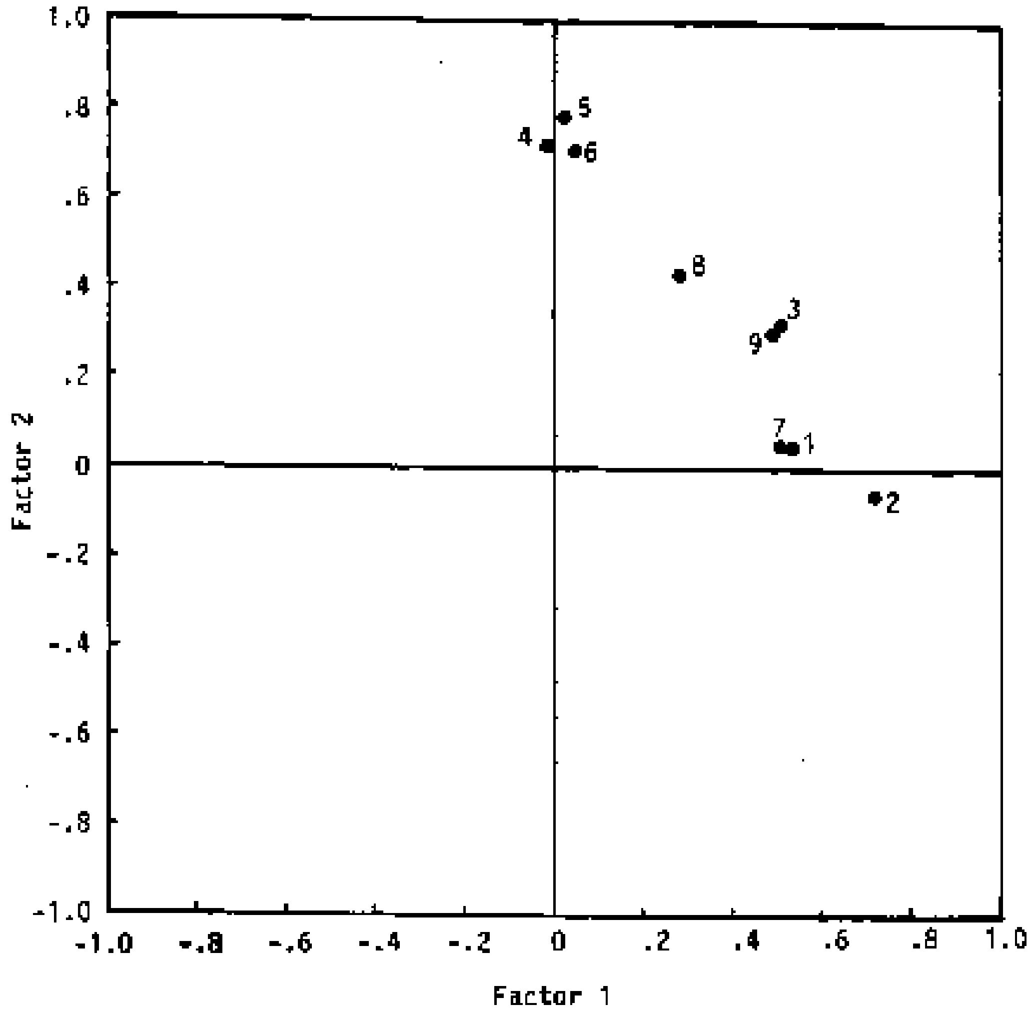


Figure 10.6a Nine Mental Tests: Rotation 1, Factor Plot for Factors 1 and 2.

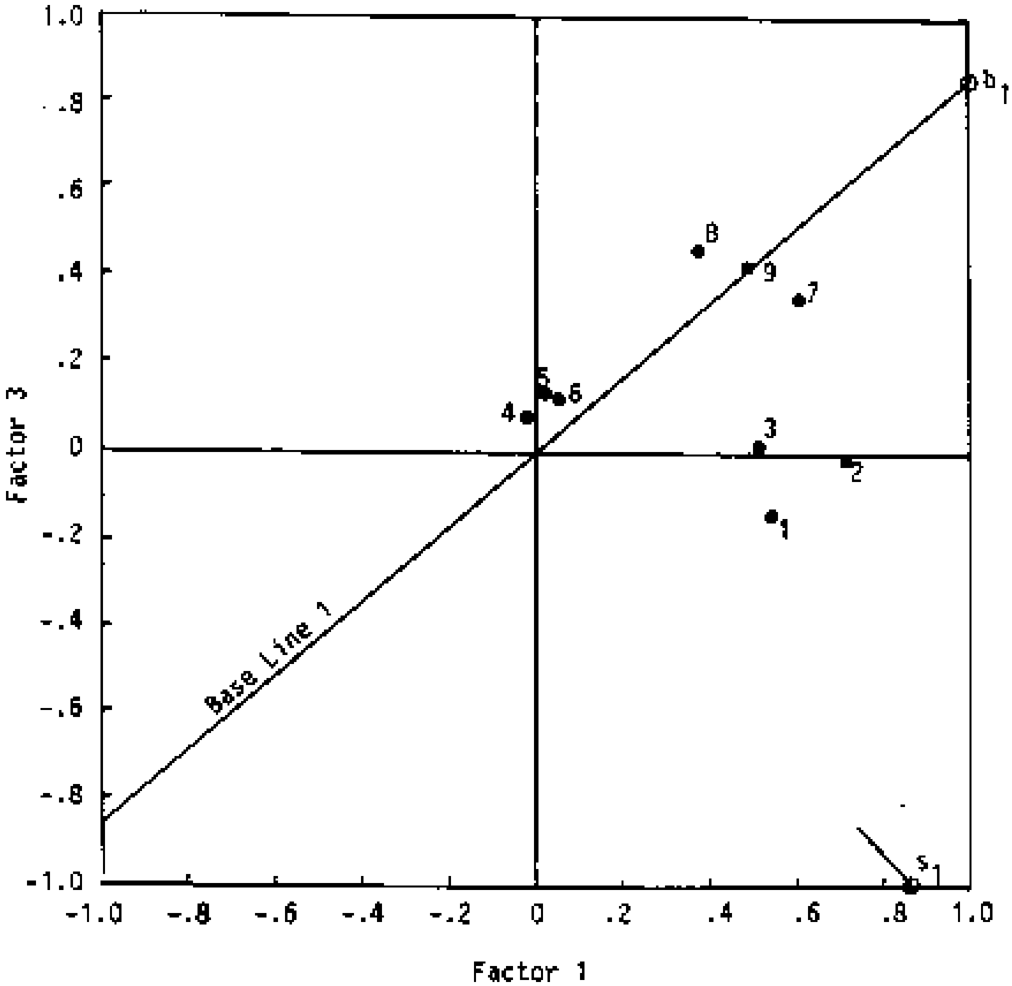


Figure 10.6b Nine Mental Tests: Rotation 1, Factor Plot for Factors 1 and 3.

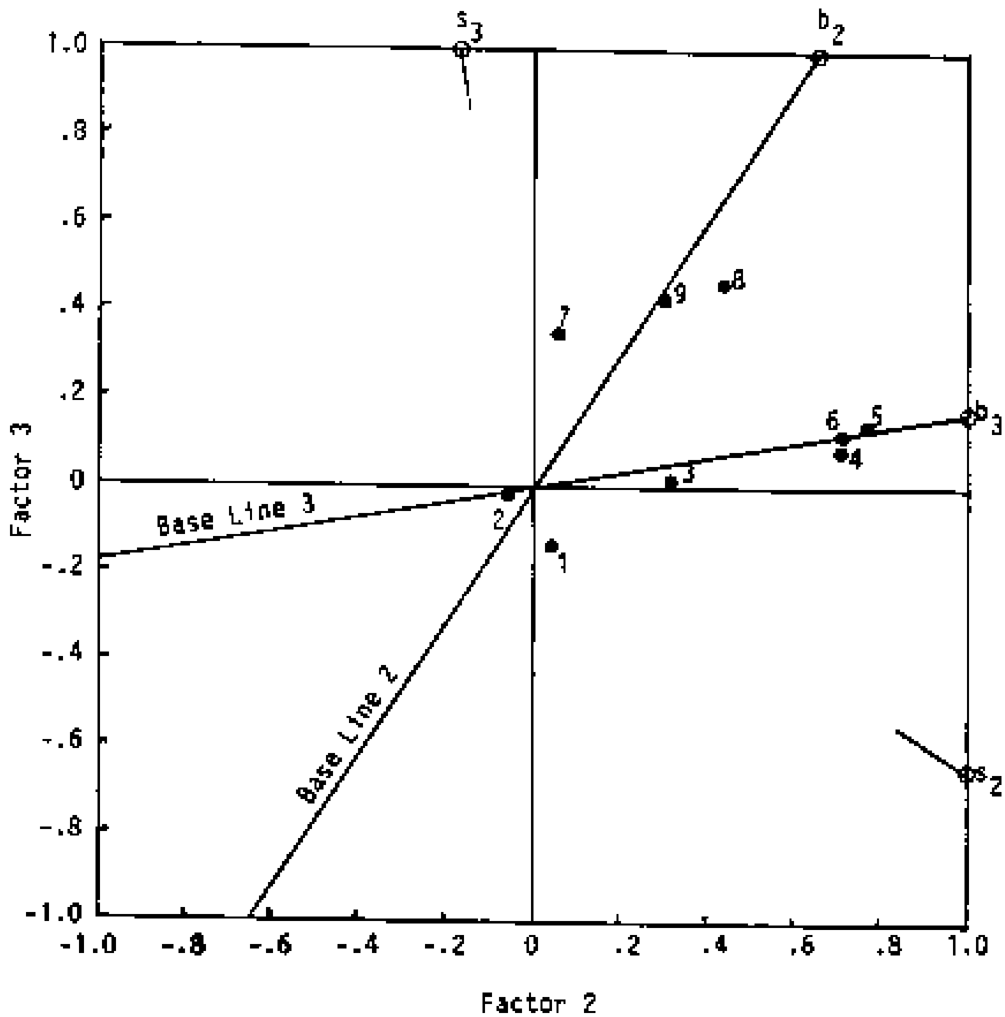


Figure 10.6c Nine Mental Tests: Rotation 1, Factor Plot for Factors 2 and 3.

Table 10.7

Graphical Transformation 2
Nine Mental tests Example

Shift Matrix S_2

	1	2	3
1	.86	.00	-1.00
2	.00	1.00	-.66
3	.00	-.17	1.00

Raw Normals Matrix $\tilde{F}_2$

	1	2	3	L^2	L
1	.232	.672	.952	1.4121	1.1883
2	.499	-.714	.629	1.1540	1.0743
3	.186	.121	-.952	.9564	.9779

Normal Matrix F_2

	1	2	3
1	.196	.565	.801
2	.464	-.665	.585
3	.190	.124	-.974

Cosines of Angles between Normals F_2F_2'

	1	2	3
1	1.00	.18	-.67
2	.18	1.00	-.56
3	-.67	-.56	1.00

Projections on Normals Matrix G_2

	1	2	3
1	.51	.12	-.15
2	.53	-.05	.00
3	.36	.29	-.05
4	-.08	.61	-.04
5	-.11	.64	.01
6	-.07	.58	.00
7	.15	-.17	.35
8	-.11	.11	.40
9	.01	.02	.38

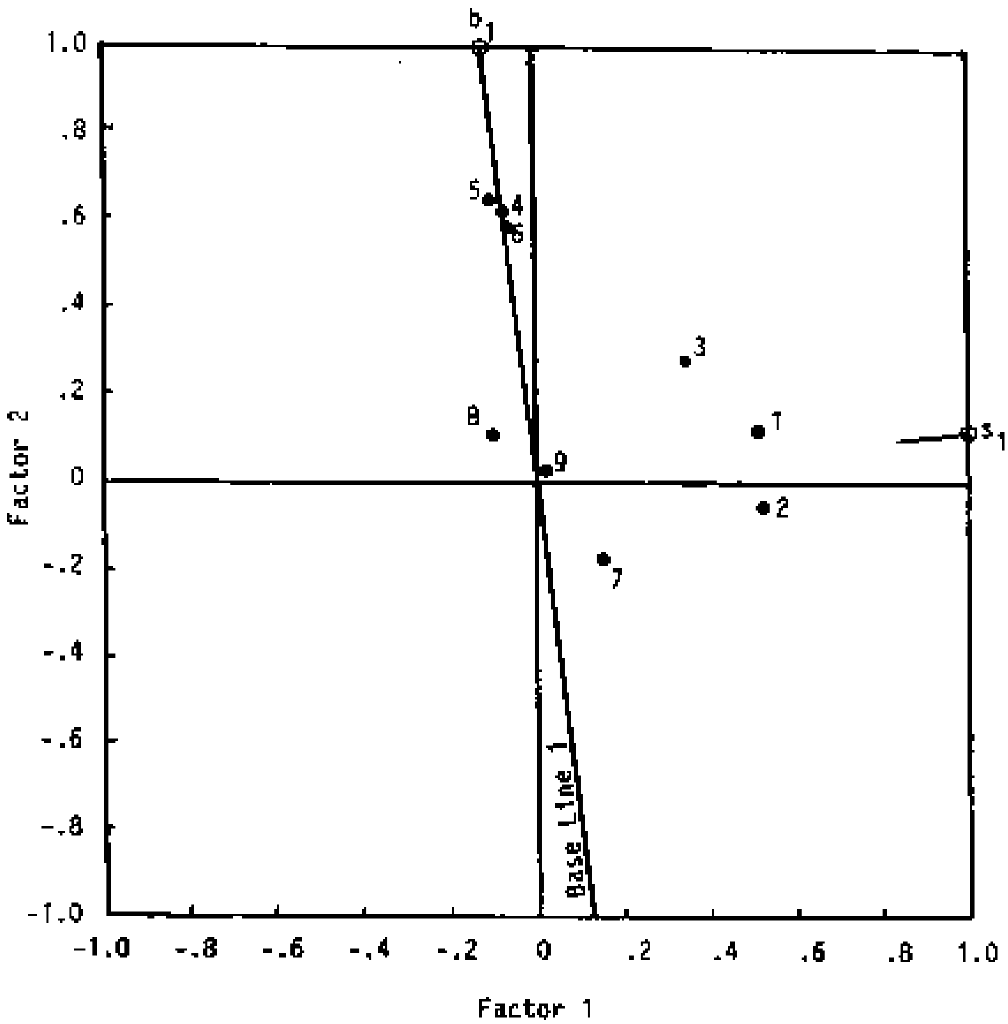


Figure 10.7a Nine Mental Tests: Rotation 2, Factor Plot for Factors 1 and 2.

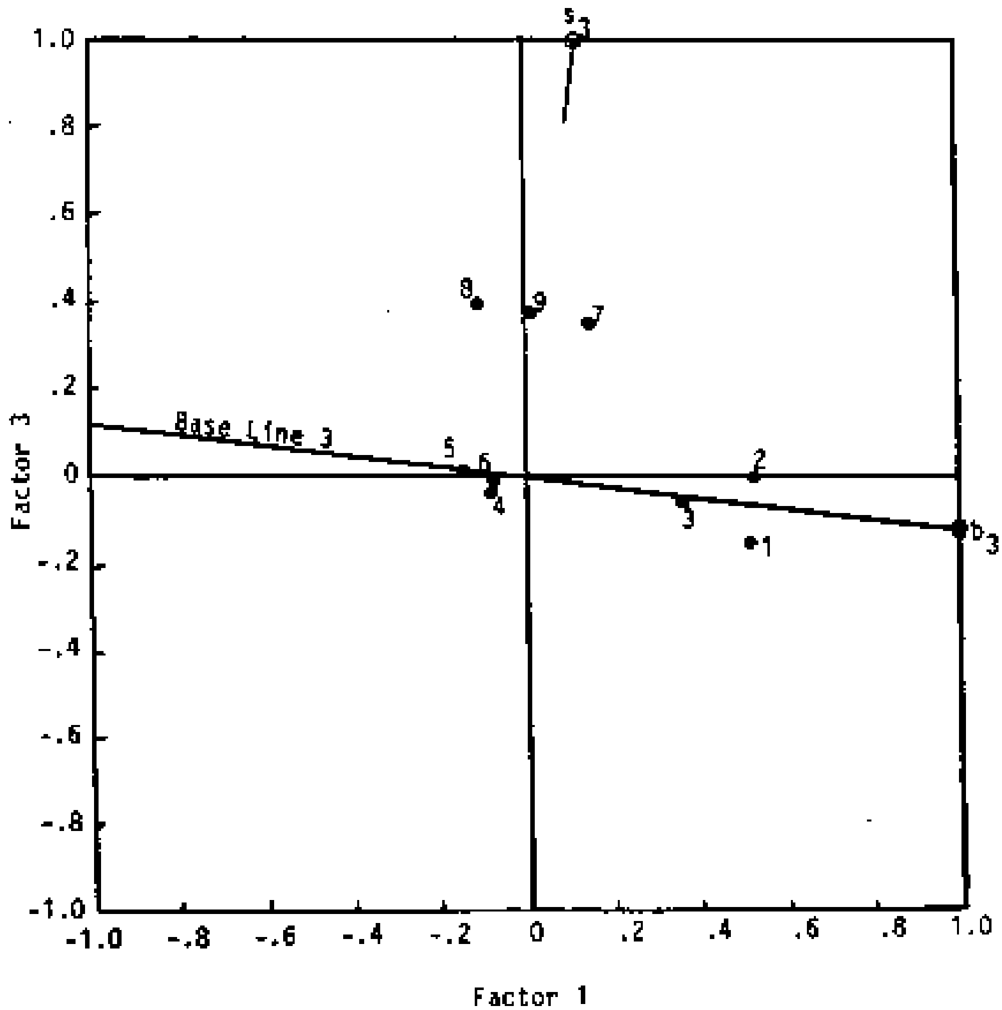


Figure 10.7b Nine Mental Tests: Rotation 2, Factor Plot for Factors 1 and 3.

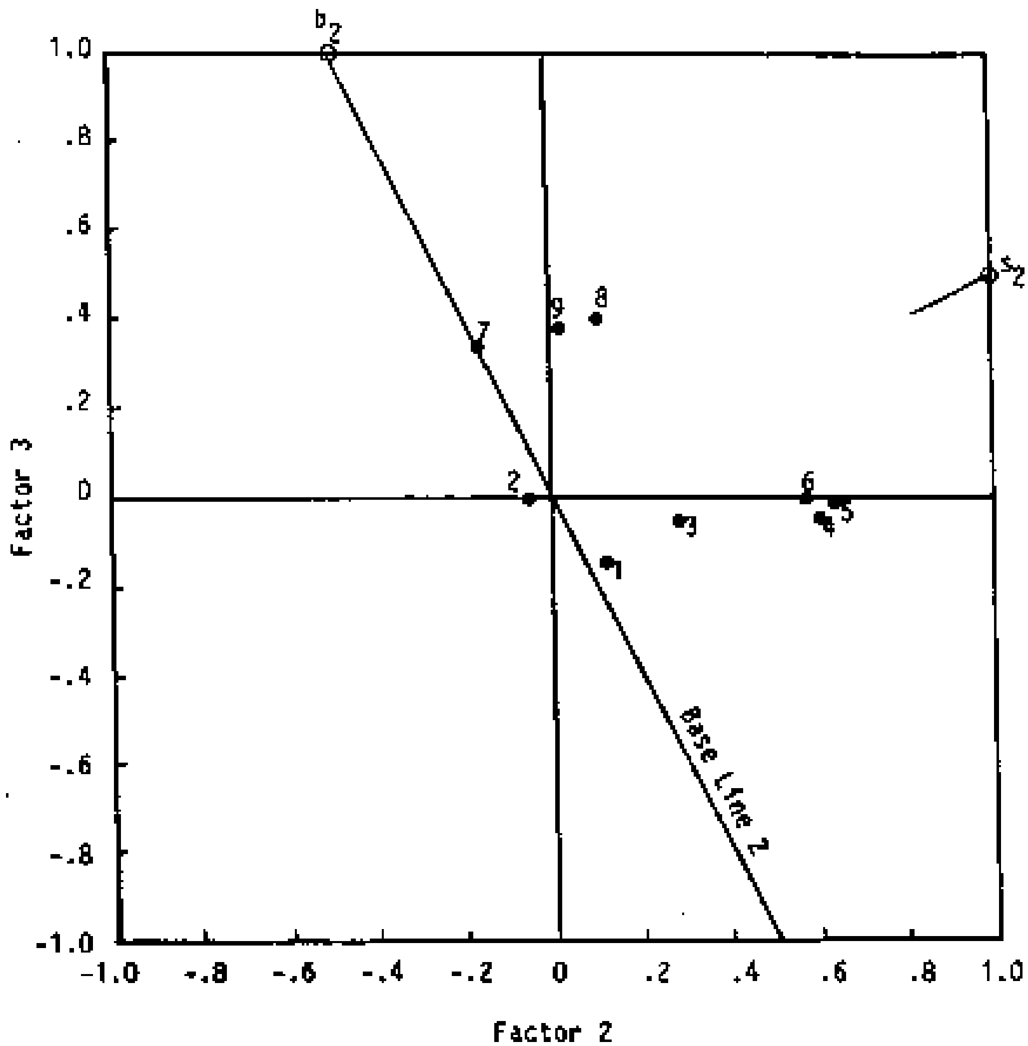


Figure 10.7c Nine Mental Tests: Rotation 2, Factor Plot for Factors 2 and 3.

Matrix $\boldsymbol{S_3}$ in Table 10.8 contains the coordinates of the shift points read from the factor plots for rotation 2 in Figures 10.7a, 10.7b, 10.7c. Raw normals matrix $\widetilde{\boldsymbol{F_3}}$ was computed by:

$$\widetilde{\boldsymbol{F_3}} = \boldsymbol{S_3 F_2}$$

and rows of $\widetilde{\boldsymbol{F_3}}$ were normalized to unit length vectors in normals matrix $\boldsymbol{F_3}$. Computations continued with matrix $\boldsymbol{F_3 F_3'}$ of cosines of angles between the normals and matrix $\boldsymbol{G_3}$ of projections on the normals.

$$\boldsymbol{G_3} = \boldsymbol{A F_3'} \ .$$

Factor plots for rotation 3 for the example are given in Figures 10.8a, 10.8b, 10.8c. These plots were inspected for further adjustments. Several quite small adjustments appeared but have not been made. Such small adjustments might be made until the analyst was completely satisfied. For the present purpose of an illustration these adjustments were not made. With this decision matrix $\boldsymbol{F_3}$ is considered to be the final normals matrix and matrix $\boldsymbol{G_3}$ is considered to be the final matrix of projections on the normals.

The general scheme of graphical transformations for rotation t is to inspect the factor plots from rotation (t $-$ 1) for adjustments to new base lines. The coordinates for these shift points are recorded as rows in matrix $\boldsymbol{S_t}$ and the new raw normals matrix $\widetilde{\boldsymbol{F_t}}$ is computed by;

$$\widetilde{\boldsymbol{F_t}} = \boldsymbol{S_t F_{(t-1)}} \ . \tag{10.14}$$

where $\boldsymbol{F_{(t-1)}}$ is the preceding rotation $\boldsymbol{F}$ matrix. Rows of $\widetilde{\boldsymbol{F_t}}$ are normalized to rows of the normals matrix $\boldsymbol{F_t}$. Projections on the new normals are in matrix $\boldsymbol{G_t}$.

$$\boldsymbol{G_t} = \boldsymbol{A F_t'} \ . \tag{10.15}$$

When no more adjustments are to be made the last normals matrix becomes the final normals matrix.

Given a final normals matrix $\boldsymbol{F}$, the complete body of transformed matrices are to be computed. Table 10.9 gives the final matrices for the example. Matrix $\boldsymbol{F}$ was given from the final graphical transformation. The other matrices were computed by the following formulas.

$$\boldsymbol{D} = [Diag(\boldsymbol{FF'})^{-1}]^{-\frac{1}{2}} \tag{10.5}$$

$$\boldsymbol{T} = \boldsymbol{D}(\boldsymbol{F'})^{-1} \tag{10.3}$$

$$\boldsymbol{T} = \boldsymbol{D}(\boldsymbol{FF'})^{-1}\boldsymbol{F} \tag{10.16}$$

$$\boldsymbol{R_{bb}} = \boldsymbol{TT'} \tag{7.1}$$

Table 10.8

Graphical Transformation 3
Nine Mental tests Example

Shift Matrix S_3

	1	2	3
1	1.00	.12	.00
2	.00	1.00	.50
3	.12	.00	1.00

Raw Normals Matrix $\tilde{F}_3$

	1	2	3	L^2	L
1	.251	.485	.872	1.0586	1.0289
2	.559	-.603	.098	.6858	.8281
3	.213	.192	-.878	.8528	.9235

Normal Matrix F_3

	1	2	3
1	.244	.472	.847
2	.675	-.728	.119
3	.231	.208	-.950

Cosines of Angles between
Normals F_3F_3'

	1	2	3
1	1.00	-.08	-.65
2	-.08	1.00	-.11
3	-.65	-.11	1.00

Projections on Normals
Matrix G_3

	1	2	3
1	.51	.06	-.09
2	.51	-.06	.07
3	.39	.32	-.01
4	-.01	.71	-.06
5	-.03	.78	.00
6	.00	.70	-.01
7	.13	.01	.39
8	-.10	.38	.42
9	.01	.25	41

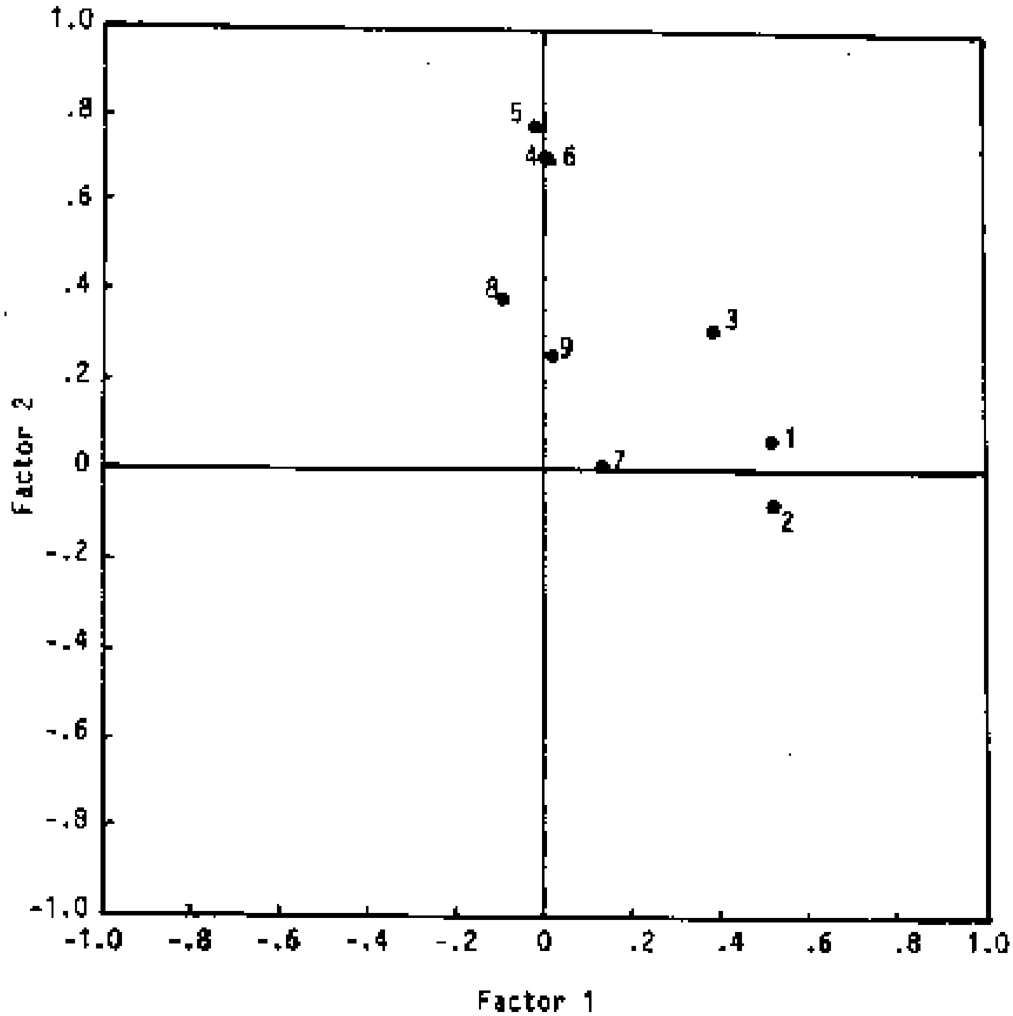


Figure 10.8a Nine Mental Tests: Rotation 3, Factor Plots for Factors 1 and 2.

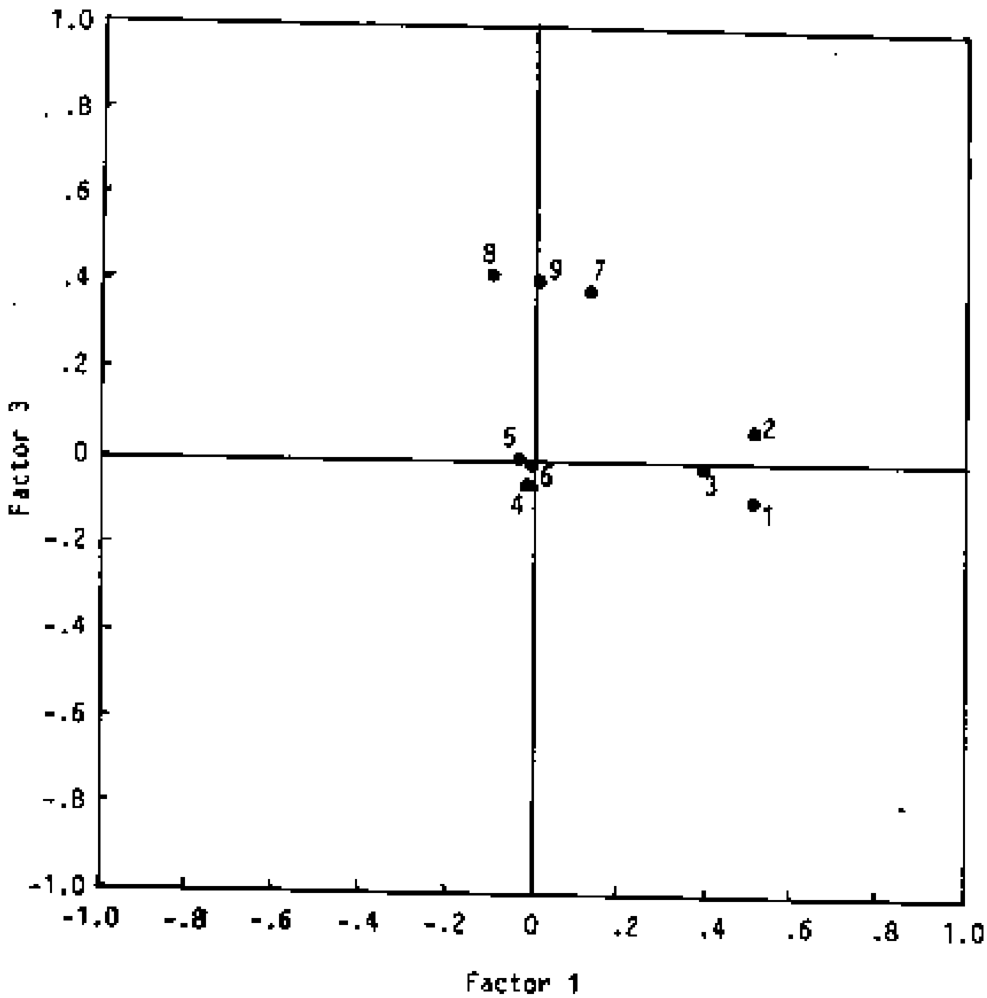


Figure 10.8b Nine Mental Tests: Rotation 3, Factor Plots for Factors 1 and 3.

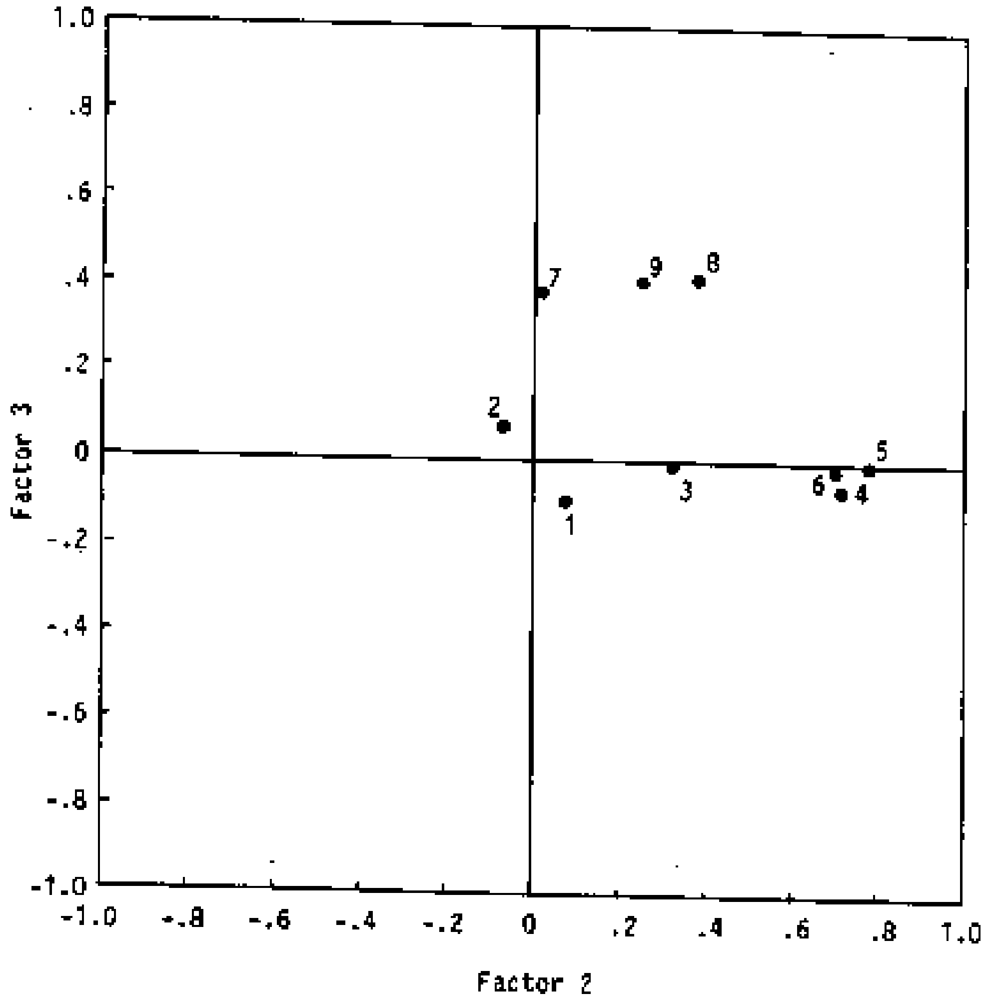


Figure 10.8c Nine Mental Tests: Rotation 3, Factor Plots for Factors 2 and 3.

Table 10.9

Graphical Transformation Matrices for Nine Mental Tests Example

Normals Matrix F

	1	2	3
1	.244	.472	.847
2	.675	-.728	.119
3	.231	.208	-.950

Traits Matrix T

	1	2	3
1	.671	.673	.310
2	.822	-.564	.077
3	.675	.545	-.498

Factor Correlation Matrix R_{bb}

	1	2	3
1	1.000	.196	.655
2	.196	1.000	.210
3	.665	.210	1.000

Cosines of Angles Between Normals and Trait Vectors

Matrix D

	1	2	3
1	.745	.000	.000
2	.000	.975	.000
3	.000	.000	.742

$$\boldsymbol{R_{bb}} = \boldsymbol{D}(\boldsymbol{F}\boldsymbol{F}')^{-1}\boldsymbol{D} \tag{10.6}$$

The final factor matrices for the example are given in Table 10.10. Formulas for these matrices are given below.

$$\boldsymbol{G} = \boldsymbol{A}\boldsymbol{F}' \tag{7.9}$$

$$\boldsymbol{B} = \boldsymbol{A}\boldsymbol{T}^{-1} \tag{7.4}$$

$$\boldsymbol{B} = \boldsymbol{G}\boldsymbol{D}^{-1} \tag{10.9}$$

$$\boldsymbol{Q} = \boldsymbol{A}\boldsymbol{T}' \tag{7.5}$$

$$\boldsymbol{Q} = \boldsymbol{B}\boldsymbol{R_{bb}} \tag{10.10}$$

Alternative formulas are given which facilitate the computations. Discussion of relations among these matrices occurs earlier in this chapter in the section for transformations for two factor studies.

10.3. Least Squares Hyperplane Fitting

While least squares hyperplane fitting involves a mathematical solution it is considered here as an adjunct to graphical transformations of factors. In order to avoid a number of small transformations of factors as well as to provide more precise definitions of the factor hyperplanes least squares solutions are obtained for projections on the normals for attribute points selected to be "in" the hyperplanes. Points "in" a hyperplane are those which have trivial projections on the normal to that hyperplane. A possible result from graphical transformations is to be able to specify which points are to be considered in each hyperplane. Then a final solution for each hyperplane can be obtained by least squares hyperplane fitting. Such a system was proposed by Tucker (1940). This method has wider applications which will not be considered here.

Before continuing with least squares hyperplane fitting a general weighted least squares proposition is required, Let $\boldsymbol{A}$, $n \times r$, be a given factor matrix on orthogonal axes. Let $\underline{\mathbf{f}}$ be a row vector with r entries which are to be determined. With $\underline{\mathbf{f}}$ as a normal vector the projections on this normal are in column vector $\mathbf{g}$ which has n entries. Vector $\mathbf{g}$ is defined by:

$$\boldsymbol{A}\underline{\mathbf{f}}' = \mathbf{g} \tag{10.17}$$

or, in expanded notation for entries g_i of $\mathbf{g}$:

Table 10.10

Graphically Transformed Factor Matrices for Nine Mental Tests Example

Matrix G of Projections on Normals

	1	2	3
1	.510	.055	-.094
2	.505	-.057	.069
3	.385	.318	-.006
4	-.009	.710	-.056
5	-.030	.781	-.002
6	.002	.704	-.006
7	.128	.008	.394
8	-.097	.378	.417
9	.008	.252	.407

Factor Weight Matrix B

	1	2	3
1	.685	.056	-.127
2	.678	-.058	.093
3	.518	.326	-.009
4	-.012	.729	-.076
5	-.041	.801	-.002
6	.002	.722	-.008
7	.172	.008	.530
8	-.130	.387	.561
9	.010	.259	.549

Matrix Q of Covariances of Modeled Attributes with Traits

	1	2	3
1	.611	.164	.340
2	.729	.094	.532
3	.576	.426	.404
4	.081	.710	.069
5	.115	.793	.139
6	.139	.721	.145
7	.526	.153	.646
8	.319	.479	.556
9	.426	.376	.610

$$\mathrm{g}_i = \sum_{j=1}^{r} a_{ij}\mathrm{f}_j \tag{10.18}$$

where a_{ij} is the ij entry in $\boldsymbol{A}$ and f_j is the j'th entry of **f̲** . Vector **f̲** is restricted being a unit vector so that:

$$\sum_{j=1}^{r} \mathrm{f}_j^2 = 1 \tag{10.19}$$

A non-negative weight, w_i , is given for each projection g_i and sum of weighted squares, designated θ , is given by:

$$\theta = \sum_{i=1}^{n} w_i \mathrm{g}_i^2 \ \ . \tag{10.20}$$

This coefficient is to be minimized. However, the restriction of equation (10.19) is to be observed. To accomplish this restricted minimum, a revised criterion is written which involves an undetermined LaGrange multiplier, β :

$$\theta = \sum_{i=1}^{n} w_i \mathrm{g}_i^2 - \beta \sum_{j=1}^{r} \mathrm{f}_j^2 \tag{10.21}$$

To minimize θ the partial derivative is taken with respect to f_j and set equal to zero.

$$\frac{\partial\theta}{\partial \mathrm{f}_j} = 2\sum_{i=1}^{n} w_i \mathrm{g}_i \frac{\partial \mathrm{g}_i}{\partial \mathrm{f}_j} - 2\beta \mathrm{f}_j = 0 \ . \tag{10.22}$$

With substitution from equation (10.18) for g_i and the partial derivative of g_i with respect to f_j equation (10.22) becomes:

$$2\sum_{j=1}^{n} w_i \sum_{k=1}^{r} a_{ik}\mathrm{f}_k a_{ij} - 2\beta \mathrm{f}_j = 0$$

or

$$\sum_{k=1}^{r} \mathrm{f}_k \sum_{i=1}^{n} w_i a_{ij} a_{ik} = \beta \mathrm{f}_j \ . \tag{10.23}$$

Define matrix $\boldsymbol{P}$ with entries p_{jk} by:

$$p_{jk} = \sum_{i=1}^{n} w_i a_{ij} a_{ik} \ \ . \tag{10.24}$$

Equation (10.23) can be written as

$$\sum_{k=1}^{r} \mathrm{f}_k p_{jk} = \beta \mathrm{f}_j$$

or, in matrix form:

$$\boldsymbol{P}\underline{\mathbf{f}}' = \beta \underline{\mathbf{f}}' \ . \tag{10.25}$$

Equation (10.25) may be written as:

$$(\boldsymbol{P} - \beta \boldsymbol{I})\underline{\mathbf{f}}' = \mathbf{0} \tag{10.26}$$

which is in the form of an eigen problem with eigenvalues β and corresponding eigenvectors $\underline{\mathbf{f}}'$. The remainder of the proposition relates β to coefficient θ .

Premultiplication of (10.24) by $\underline{\mathbf{f}}$ yields:

$$\underline{\mathbf{f}}\boldsymbol{P}\underline{\mathbf{f}}' = \beta \underline{\mathbf{f}}\underline{\mathbf{f}}'$$

which with the restriction of (10.19) yields:

$$\beta = \underline{\mathbf{f}}\boldsymbol{P}\underline{\mathbf{f}}' \tag{10.27}$$

From (10.20) and equation (10.18) coefficient θ may be written as:

$$\theta = \sum_{i=1}^{n} w_i \sum_{j=1}^{r} a_{ij} \mathrm{f}_j \sum_{k=1}^{r} a_{ik} \mathrm{f}_k = \sum_{j=1}^{r} \mathrm{f}_\mathrm{j} \sum_{\mathrm{k}=1}^{r} \mathrm{f}_\mathrm{k} \sum_{\mathrm{i}=1}^{\mathrm{n}} a_{\mathrm{ij}} w_i a_{ik}$$

which with equation (10.24) yields:

$$\theta = \sum_{j=1}^{r} \mathrm{f}_j \sum_{k=1}^{r} \mathrm{f}_k p_{jk}$$

or, in matrix form:

$$\theta = \underline{\mathbf{f}}\boldsymbol{P}\underline{\mathbf{f}}' \ .$$

Comparison with equation (10.27) yields:

$$\theta = \beta \ .$$

Consequently, minimum θ is obtained by using as f' the eigenvector corresponding to the least eigenvalue. Minimum θ is the value of the least eigenvalue.

The preceding development is stated in a more powerful form than will be used in least squares hyperplane fitting in that there is no restriction on the weights other than they be nonnegative. In least squares hyperplane fitting the weights are restricted being either 0 or 1. In other contexts to be considered in the next chapter these weights will have other positive values.

Least squares hyperplane fitting is a single plane method in that it operates on each transformed factor individually. A separate solution is required for each of the factors in a study. The computations for a factor start from a selection vector of weights for the attributes these weights being restricted to values of 0 or 1 . A weight of 0 is assigned to each attribute judged to be "out of the hyperplane", that is, to have a significant projection on the derived normal. A weight of 1 is assigned to each attribute judged to be "in the hyperplane", that is, to have a trivial projection on the derived normal. Consider Table 10.11 which gives the results for the nine mental tests example. The first column of the selection vectors section gives the weights for the first factor. These selections are based on the final graphical transformation. Attributes 1, 2, 3, and 7 were judged to be "out" of the hyperplane so that they were given weights of 0. The remaining attributes were judged to be "in" the hyperplane and were given weights of 1. Note that these values are weights used in hyperplane fitting. An easy misconception is that this selection vector represents an anticipated pattern of loadings with l's being interpreted as high and 0's being interpreted as low. The reverse of this misconception is true. 0's represent attributes which are to have high valued projections and l's represent attributes which are to have low valued projections. 0 is for "out; 1 is for "in".

Column 1 of the remaining three sections of Table 10.11 give results for the first factor. There are three eigenvalues of the product matrix with a high first eigenvalue, a moderate second eigenvalue and a small third eigenvalue, this third eigenvalue being the sum of squared projections for the attributes "in" the plane. These results indicate a good fit of the plane for the selected attributes. Column 1 of the normals matrix contains the third eigenvector for the first factor product matrix. Projections on this normal are given in column 1 of the bottom matrix. Inspection of this first column of projections indicates a satisfactory solution. Attributes judged to be "in" the plane in the selection vector have projections in the interval bounded by $\pm$.10. Attributes judged to be "out" of the plane have projections greater than .10. Note, also, that all attributes "out" of the plane have positive projections. This satisfies the principle of a positive manifold.

Columns 2 and 3 of all sections of Table 10.11 give the results for factors 2 and 3 of the example. These results, also, appear to be satisfactory. These solutions provide a normals matrix $\boldsymbol{F'}$. Note that it is in transpose form. From this matrix the transformed factor solution can be computed as was done in the last part of the preceding section.

Two classes of problems may occur in least squares hyperplane fitting. First, not all attributes selected to be "in" a hyperplane end up with projections outside the limits for being in the hyperplane. This condition can be detected by the last eigenvalue not being small enough.

Table 10.11

Least Squares Hyperplane fitting
Nine Mental Tests Example

Selection Vectors
Entries used as Weights

	1	2	3
1	0	1	1
2	0	1	1
3	0	0	1
4	1	0	1
5	1	0	1
6	1	0	1
7	0	1	0
8	1	0	0
9	1	0	0

Eigenvalues of Product Matrix

	1	2	3
1	2.101	1.220	1.979
2	.446	.131	1.023
3	.006	.006	.015

Normals matrix F Transposed

	1	2	3
1	.295	.684	.255
2	.495	-.725	.187
3	.816	.084	-.949

Projections on Normals

	1	2	3
1	.531	.050	-.091
2	.537	-.057	.069
3	.415	.317	.005
4	.005	.710	-.034
5	-.011	.783	.023
6	.020	.706	.017
7	.169	.021	.397
8	-.057	.393	.431
9	.050	.267	.418

Also, the projections on the normal can be inspected for such points not within bounds. The second class of problems involves degenerate selection of attributes to be "in" the hyperplane. The selected points may be in a space of two or more dimensions. The hyperplane is not well defined by the selection of points. In this case two or more eigenvalues of the product matrix will be extremely small. Note that each eigenvalue is the sum of squared projections on the corresponding eigenvector. To illustrate this case a degenerate selection vector was written for the nine mental tests example. This selection vector is given in Table 10.12 along with associated computations. Note that the last two eigenvalues are very small which indicates that the selected attributes, 4, 5, and 6, form a cluster in the space of this example. All three eigenvectors and projections on these eigenvector are given. Note that the three selected attributes have projections in the range of $\pm.10$ on both of the last two eigenvectors. Figure 10.9 presents a plot of the projections on these last two eigenvectors. Points 4, 5, and 6, form a tight cluster at the origin. To eliminate such a degeneracy one or more points should be added to the selection of points to be "in" the hyperplane. In the example from Figure 10.9 a judgment might be made to add points 8 and 9. Such a move would result in the selection vector for factor 1 in Table 10.11. An alternative would be to add points 1, 2, and 3 which would result in the selection vector for factor 3 of the example as given in Table 10.11. Either of these selections would eliminate the degeneracy.

Selection vectors can be developed by other means than as a result from graphical transformations. In a tradition of confirmatory studies selection vectors may be hypothesized from previous studies and results studied toward confirmation of these hypotheses. Care must be taken in this class of use to guard against the problems described in the preceding paragraph. Another alternative is to develop selection vectors from other methods of factor transformation. One particular suggestion was by Kaiser and Cerny (1978) in an article entitled "Casey's method for fitting hyperplanes from an intermediate orthomax solution." In this method they started from an orthomax solution such as a VARIMAX transformation (to be described in the next chapter) and selected attributes to be "in" hyperplanes in terms of loadings being less in absolute values than computed critical values. They suggested a method for computing these critical values. Since this method has not stood up well in Monte Carlo studies, it will not be discussed further here.

Table 10.12

Degenerate Example of LSQHYP Fitting
Nine Mental tests Example

Selection Vectors

1	0
2	0
3	0
4	1
5	1
6	1
7	0
8	1
9	1

Eigenvalues of Product Matrix

1	1.657
2	.003
3	.000

Eigenvectors of Product Matrix

	1	2	3
1	.790	.561	.245
2	-.606	.659	.446
3	.089	-.501	.861

Projections on Eigenvectors

	1	2	3
1	.139	.333	.505
2	.058	.539	.494
3	.402	.353	.385
4	.714	-.045	.005
5	.793	.012	-.017
6	.720	.031	.013
7	.118	.638	.115
8	.456	.488	-.100
9	.347	.562	.001

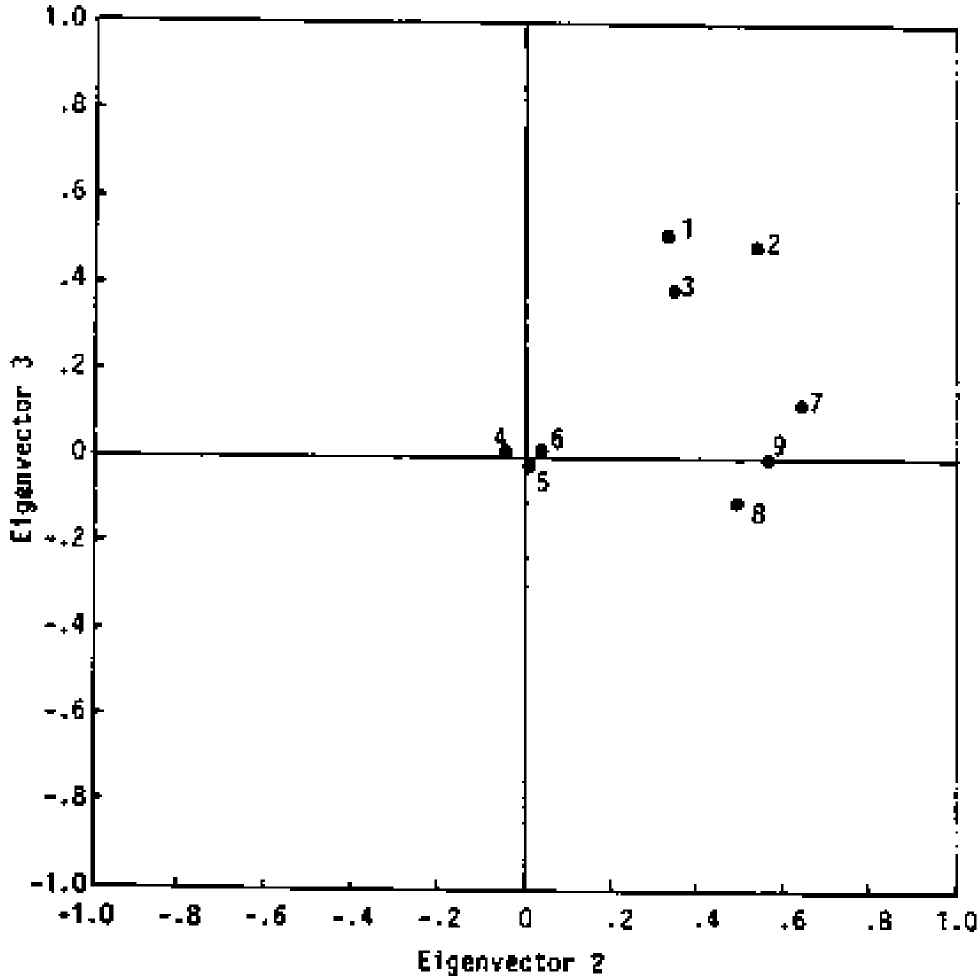


Figure 10.9 Plot of attribute point projections on eigenvectors 2 and 3 for degenerate LSQHYP fitting example, nine mental tests example.

REFERENCES FOR CHAPTER 10

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

Thurstone, L. L. & Thurstone, Thelma Gwinn. Factorial studies of intelligence. Psychometric Monograph No. 2. Chicago: University of Chicago Press, 1941.

Tucker, L. R. A rotational method based upon the mean principal axis of a sub-group of tests. Psychological Bulletin, 1940, 37, 578.

Kaiser, Henry F. & Cerny, Barbara A. Casey's method for fitting hyperplanes from an intermediate orthomax solution. Multivariate Behavioral Research, 1978, 13, 395-401.