CHAPTER 12
FACTOR OPERATIONS FOR MULTIPLE DATA

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

CHAPTER 12
FACTOR OPERATIONS FOR MULTIPLE DATA

A number of situations exist which involve operations with two or more bodies of data. This is in contrast to the preceding considerations of operations for a single body of data. The sections in this chapter consider such situations as when after a factor analysis has been performed for one battery of attributes data is available for additional attributes such as course grades following the analysis of a battery of pre-course aptitudes data. A second example involves transformations of an original factor matrix to hypothetical loadings. Also to be considered are analyses and tranformations for covariance matrices and factor matrices from several samples.

Scales of measurement of attribute measures and of factors are important mattesr when dealing with data from several samples. The common factor model in the sample was discussed in Chapter 4. Equation (4.19) indicates that population factor weights are multiplied by sample factor standard deviations to obtain the sample factor weights. These factor weights are for analyses of covariance matrices. When analyses are made of sample correlation matrices the factor weights for each attribute are divided by the sample standard deviation for that attribute as indicated in equation (4.34). Thus, there are column wisea and406 row wise constants of proportionality for the factor loadings when analyses are made of correlation matrices. These scaling constants must be considered in the analyses involving multiple samples.

Analyses in several populations present extended problems as compared with analyses in samples from a single population. This topic is relevant to several of the sections in this chapter and especially to the section on factor analyes in several populations. A most important point is that the attribute measures should be expressed in a common scale across the populations. For a small example see Table 12.1 and Table 12.2 which involves measures on five attributes for a sample of first-year medical students and a sample of forth-year medical students. In that differential selection and different experiences have occured for the students in these two samples they may be considered to have been drawn from two populations. For the first three tests the fourth-year students scored lower and had markedly higher standard deviations than did the first-year students. There is a need to pool the data for these two samples to obtain some type of standardization similar to converting each covariance matrix to a correlation matrix. One possibility is to obtain the unweighted mean covariance matrix to obtain pooled variancse from which pooled standard deviations may be obtained. Some individuals would like to weight each covariance matrix by the sample size in obtaining a mean. This would be reasonable if the two samples are considered to be drawn from the same population. However, when there are two

Table 12.1

Multiple Populations Sample Means and Standard Deviations
Medical Students Data*

		First-Year Students (N = 141)		Fourth-Year Students (N = 75)	
	Tests	Mean	SD	Mean	SD
1	Logical Diagrams	10.8	2.68	10.4	2.95
2	Locations	9.9	2.74	8.7	3.43
3	Resourceful Arithmetic	21.5	3.02	19.3	4.30
4	Vocabulary	23.0	5.70	24.5	5.80
5	Word Meaning	16.0	4.79	16.3	4.61

*Data on cognitive process tests in study by Frederiksen, et al, (1980) on “Development of methods for selection and evaluation in undergraduate medical education.”

Table 12.2
Multiple Populations Sample covariance matrices
Medical Students Data

Raw Covariance Matrices | Scaled Covariance Matrices

First –Year Students, N = 141

Raw Covariance Matrices

	1	2	3	4	5
1	7.2	2.5	2.4	1.3	.8
2	2.5	7.5	1.9	-2.0	-1.3
3	2.4	1.9	9.1	-.5	.3
4	1.3	-2.0	-.5	32.5	20.4
5	.8	-1.3	.3	20.4	22.0

Scaled Covariance Matrices

	1	2	3	4	5
1	.91	.29	.22	.08	.06
2	.29	.78	.16	-.11	-.09
3	.22	.16	.66	-.02	.02
4	.08	-.11	-.02	.98	.76
5	.06	-.09	.02	.76	1.02

Fourth –Year Students, N = 141

Raw Covariance Matrices

	1	2	3	4	5
1	8.7	3.2	3.8	5.4	3.9
2	3.2	11.8	6.1	3.7	3.0
3	3.8	6.1	18.5	5.3	4.2
4	5.4	3.7	5.3	33.6	20.8
5	3.9	3.0	4.2	20.8	21.2

Scaled Covariance Matrices

	1	2	3	4	5
1	1.09	.37	.36	.33	.30
2	.37	1.22	.53	.21	.21
3	.36	.53	1.34	.25	.24
4	.33	.21	.25	1.02	.78
5	.30	.21	.24	.78	.98

Mean Covariance Matrices

Raw Covariance Matrices

	1	2	3	4	5
1	8.0	2.9	3.1	3.4	2.4
2	2.9	9.6	4.0	.8	.8
3	3.1	4.0	13.8	2.4	2.2
4	3.4	.8	2.4	33.0	20.6
5	2.4	.8	2.2	20.6	21.6

Scaled Covariance Matrices

	1	2	3	4	5
1	1.00	.33	.29	.21	.18
2	.33	1.00	.34	.05	.06
3	.29	.34	1.00	.11	.13
4	.21	.05	.11	1.00	.77
5	.18	06	.13	.77	1.00

different populations there is a question as to what covariance matrix is being estimated. There would be a desirability not to over weight a larger sample so as to overshadow relations appropriate for the population from which the smaller sample was drawn. The reciprocals of the pooled standard deviations are used as row and column multipliers of the sample covariance matrices to obtain the scaled covariance matrics given in Table 12.2 . Note that the diagonals of the mean scaled covariance matrices are unity.

12.1 Factor Extension Techniques

A number of factor analytic studies involve a major battery of attributes followed by some supplementary attributes to which the analysis of the major battery is to be extended. For an example consider a battery of aptitude measures followed by measures of some criteria. There is a desire to extend the analysis of the aptitude battery to the criteria measures so as to find the dependence of the criteria on the aptitude factors. Data for a small example are based on a study conducted by French, et al, (1952) on entrance tests and course grades at the United States Coast Guard Accademy. The correlations given in Table 12.3 are based on notes by Tucker from his participation in this study. There had been extensive selection of the cadets in the class based on entrance tests; however, the aptitude tests used in the illustation were not used in these selections. Never the less, the statistics for the tests in the example may have been influenced by incidental selection, a point to consider during inspection of results of the example. Following the five aptitude measures in the example are course grades in four first term courses. The purpose in our example is to relate the course grades to the factors from analysis of the aptitude measures. A second type of example for factor extension involves a battery of aptitude measures followed by some personality measures. In this type of example the purpose is to investigate the dependence of the personality measures on the aptitude factors. There are a number of other types of examples. The main purpose of factor extension is to relate supplementary measures to factors determined from a major battery of measures.

Chapter 7 provides the basic framework for factor fitting as given in equation (7.16). For the present purposes this equation may be written as:

$$\Delta = (C - U^2) - AA' \qquad (12.1)$$

where Δ is a matrix of residuals. With the two batteries matrix C is partitioned into the four sections illustrated in Table 12.3. Matrix Δ is similarly partitioned. Furthter, there are sections of the factor matrix A : A_1 and A_2 for the two batteries. These partitionings result in the following equations:

$$\Delta_{11} = (C_{11} - U_1^2) - A_1A_2' \quad ; \qquad (12.2.11)$$

$$\Delta_{12} = C_{12} - A_1A_2' \quad ; \qquad (12.2.12)$$

Table 12.3
Illustration of Factor Extension Procedure
Correlation Matrix*

Section C_{11}

	1	2	3	4	5
1	1.00	.59	.31	.22	.14
2	.59	1.00	.16	.10	.03
3	.31	.16	1.00	.47	.39
4	.22	.10	.47	1.00	.47
5	.14	.03	.39	.47	1.00

Section C_{12}

	6	7	8	9
1	.25	.22	.09	.09
2	.38	.33	.11	.10
3	-.05	.06	.32	.35
4	-.10	.06	.31	.34
5	-.05	.02	.35	.41

Section C_{21}

	1	2	3	4	5
6	.25	.38	-.05	-.10	-.05
7	.22	.33	.06	.06	.02
8	.09	.11	.32	.31	.35
9	.09	.10	.35	.34	.41

Section C_{22}

	6	7	8	9
6	1.00	.56	.31	.24
7	.56	1.00	.42	.36
8	.31	.42	1.00	.76
9	.24	.36	.76	1.00

* These correlations are based on the study by French, et al, (1952) on entrance tests and course grades at he United States Coast Guard Academy. The attributes are:

1. American Council on Education Test, Linguistics Section.
2. College Entrance Examination Board, SAT – Verbal.
3. College Entrance Examination Board, Secondary School Physics.
4. American Council on Education Test, Quantitative Section.
5. College Entrance Examination Board, SAT – Mathematics.
6. First Term Grade in History and Literature.
7. First Term Grade in English Composition.
8. First Term Grade in Physics.
9. First Term Grade in Analytic Geometry and Calculus.

$$\Delta_{21} = C_{21} - A_2A_1' = \Delta_{12}' \quad (12.2.21)$$

$$\Delta_{22} = (C_{22} - U_2^2) - A_2A_2' \quad .; \quad (12.2.22)$$

Diagonal matrices U_1^2 and U_2^2 are the uniqueness for batteries 1 and 2. Note that the off-diagonal sections do not involve the uniqueness in this formal model; however, some cases with real data, this model may have to be adjusted to include the uniqueness of some attributes in batteries 1 and 2. These cases will be discussed subsequaently.

Factor extension involves a least squares solution for the entries in matrix Δ_{21} (equivalently in matrix Δ_{12}) which results, according general least squares solutions, in:

$$A_2 = C_{21}A_1(A_1'A_1)^{-1} \quad . \quad (12.3)$$

When this formula is applied to a body of data the results should be inspected for possible violations of the assumptions implied by equations (12.2). Interpretations of these results may be seriously effected by these violations of assumptions. Possible such problematic results will be discussed in terms of the example using the Coast Guard Accademy data.

As previously indicated, the correlation matrix for the Coast Guard Accademy data is given in Table 12.3. This matrix is sectioned into matrices C_{11} , C_{21} , C_{12} , C_{22} . A principal factors solution of C_{11} using the Jöreskog and Goldberger (1972) initial uniqueness (see Chapter 9, equation 9.97) is given as matrix A_1 in Table 12.4 . Extension matrix A_2 was obtained by equation (12.3). The residual matrix with sections given in equation (12.2.11 through (12.2.22) is given in Table 12.5. These matrices are to be inspected for violations of the assumptions of the factor extension model. Consider section Δ_{21} , these residuals appear quite small indicating a good application of the extension model. However, the column of negative residuals for attribute 1 with grades 6 - 9 along with the column of positive residuals for attribute 2 with theses grades and the column of positive residuals for attribute 5 indicates some small general effect which may have arisen from the selective process for this class. Such an effect might be noted in the interpretation of the results. An effect which might have been anticipated did not occur. This effect was a possible high residual between the two measures of physics performance: attribute 3, Secondary School Physics, and attribute 8, Forst Term Grade in Physics. If this residual had been high, the interpretation would be that there was an overlaping specific between the two attributes. Section Δ_{22} has high residuals among the grades attributes which indicates a further factor among these attributes which does not overlap with the aptitude battery. This appears to be the 'grades factor' found in the analysis by French, et al, (1952). A factor extracted from section Δ_{22} would be appended to matrix A_2 with zero loadings for matrix A_1 .

Results of a factor transformation of matrix A_1 by the DAPPFR method are given in

Table 12.4
Illustration of Factor Extension Procedure
Original Factor Matrices

Factor Matrix A_1*

	1	2
1	.61	.45
2	.47	.56
3	.63	-.21
4	.61	-.34
5	.52	-.39

Factor Matrix A_2

	1	2
6	.13	.47
7	.23	.30
8	.42	-.25
9	.46	-.30

*Principal factors of C_{11} using the Joreskog and Goldberger (1972) initial uniqueness (see Chapter 9, equation 9.97).

Table 12.5
Illustration of Factor Extension Procedure
Residual Matrix

Section Δ_{11}

	1	2	3	4	5
1	.42	.05	.02	.00	.00
2	.05	.47	-.02	.00	.00
3	.02	-.02	.56	.01	-.02
4	.00	.00	.01	.51	.02
5	.00	.00	-.02	.02	.58

Section Δ_{12}

	6	7	8	9
1	-.04	-.05	-.05	-.05
2	.06	.06	.05	.06
3	-.03	-.02	.01	.00
4	-.02	.02	-.03	-.04
5	.07	.02	.04	.06

Section Δ_{21}

	1	2	3	4	5
6	-.04	.06	-.03	-.02	.07
7	-.05	.06	-.02	.02	.02
8	-.05	.05	.01	-.03	.04
9	-.05	.06	.00	-.04	.06

Section Δ_{22}

	6	7	8	9
6	.76	.39	.37	.32
7	.39	.86	.40	.35
8	.37	.40	.77	.50
9	.32	.35	.50	.70

factor extracted from section Δ_{22} would be appended to matrix A_2 with zero loadings for matrix A_1 .

Results of a factor transformation of matrix A_1 by the DAPPFR method are given in Table 12.6. The structure transformation results are given at the left with the pattern transformations results being given at the right. This transformation is carried on to matrix A_2 . Factor 1 appears to be a linguistics factor while factor 2 appears to be a quantitative factor. These two factors carry over nicely from the aptitude battery to the grades battery. There is one surprize: the large negative loading for attribute 5, First Term Grades in History and Literature. The interpretation of this negative loading would be a challenge to the analyst. It will be left as an observation in the present context.

A possible extension of equation (12.3) would results in matrix G_2 directly from matrix C_{21} using matrices G_1 and product matrix FF' . This will not be pursued here; the simpler solution appears to apply the extension procedure to obtain matrix A_2 followed by the factor transformation used in the example. A similar possible extension of equation (12.3) would results in matrix B_2 directly from matrix C_{21} using matrices B_1 and product matrix TT' . Again, this extension will not be pursued here.

12.2 Factor Transformations to an Hypothesized Matrix of Loadings: Procrustes Transformations

The class of situations considered in this section concerns the transformation of an original factor matrix A to approximate hypothetical loadings defined by the analyst. These hypothetical loadings are in a matrix H . The analyst may specify the hypothetical loadings from previous experience such as the analysis of the battery of attributes in a previous study. Alternatively, the analyst may define the hypothetical loadings in terms of some theory. In a sense these operations verge on confirmatory analysis in which the accuracy of the hypothetical loadings is checked. A variety of techniques will be considered in this section. A major division between techniques concerns whether or not the transformed factors are restricted to being uncorrelated. A first subsection will be concerned with oblique factor transformations and a second subsection will consider orthogonal factor transformations.

Following is an outline of general notoation to be employed and conditions to be considered.

A is original, uncorrelated factor matrix with: n attributes ; r factors.

H is matrix of hypothetical factor loadings. There are to be n rows in H ; however, in cases (to be noted) there may be fewer than r columns of H . Let $\underline{h}_m$ be the m'th column of H .

Table 12.6
Illustration of Factor Extension Procedure
Factor Transformation*

Structure Transformation Normals, F'			Pattern Transformation Traits, T'		
	1	2		1	2
1	.47	.69	1	.72	.88
2	.88	-.72	2	.69	-.47

Transformed Factor Loadings

Projections on Normals G_1			Factor Weights B_1		
	1	2		1	2
1	.69	.09	1	.73	.10
2	.71	-.08	2	.75	-.09
3	.12	.59	3	.12	.62
4	-.01	.67	4	-.01	.70
5	-.09	.64	5	-.10	.67

G_2			B_2		
	1	2		1	2
6	.48	-.25	6	.50	-.27
7	.37	-.06	7	.39	-.06
8	-.02	.47	8	-.02	.49
9	-.05	.53	9	-.05	.56

Product Matrices

FF'			Φ = TT'		
	1	2		1	2
1	1.00	-.31	1	1.00	.31
2	-.31	1.00	2	.31	1.00

* DAPPFR transformation of A_1.

Some or all entries in each $\underline{h}_m$ may be specified by the analyst. Let I_m be the set of attributes, i , for which the entries in $\underline{h}_m$ are specified. A sum over attributes in I_m is symbolized by: $\sum_{i \in I_m}$. A matrix $\tilde{A}_m$ is obtained by deleting from A rows for which the entries in $\underline{h}_m$ are not specified. Let $\tilde{\underline{h}}_m$ be the column of $\underline{h}_m$ with the unspecified entries deleted.

W is a matrix of weights to be defined to minimize a given least squares criterion. There is to be a column, $\underline{w}_m$, for each column of $\underline{h}_m$.

V is a matrix of derived factor loadings with columns $\underline{v}_m$.

$$AW = V \quad ; \tag{12.4M}$$

$$A\underline{w}_m = \underline{v}_m \quad . \tag{12.4V}$$

Entries in V are to be matched with entries in H in a least squares sense according to some chosen criterion.

12.2.1 Oblique Factor Transformations

A first issue for oblique factor transformations concerns the type of factor loadings. There are three cases. First is the case where there is no given restriction on the weights. This is the raw weights case. Second is the case where each weight vector is restricted to unit length so as to be the normal to an hyperplane with the factor loadings being projections on the normal (structure loadings). This is the normals case. Third is the case where each column of V is to be a column of T^{-1} so that the factor loadings are factor weights in matrix B (pattern loadings). This is the trait vectors case.

For the first two cases there may be fewer than r columns in H so that this matrix is incomplete. Any remaining columns of W and V are to be defined separately so as to provide a complete transformation. For the third case of factor weights (pattern loadings) matrix H must be complete in the sense of having r columns.

A second issue is whether the entries in H are to be taken in a fixed sense or are to be taken as known only within an undefined constant of proportionality for each column of H . This idea of undefined proportionality is supported by two matters discussed below.

(1) To permit freedom to the analyst to utilize hypothesized loadings defined within a constant of proportionality and not require the analyst to define these loadings at an exact level.

(2) To take into account the effects of sampling described in Chapter 4 whereby sample factor weights are proportional by columns to population weights.

An important point related to sampling sffects discussed in Chapter 4 is that in several samples the measures for each attribute should be measured on a common scale across the samples. This may be accomplished, instead of using a correlation matrix for each sample, by standardizing the covariance matrix for each sample using a combined standard deviation.

An undefined constant of proportionality for each factor is symbolized as c_m . For each of the weights vector cases there are two hypothetical loadings cases: the fixed hypothetical loadings case and the proportional hypothetical loadings case.

A 3 x 2 type design of matching criteria are considered in this note: three types of loadings by two conditions whether or not proportionality of columns of H are considered. In addition, a canonical type criterion is considered.

Operations for the oblique factor transformations will be illustrated using the Medical Students Data described earlier. The scaled covariance matrices given in Table 12.2 are used to meet the joint standardization requirement described earlier in this section. Table 12.7 gives the principal factor matrices for the two samples. Hypothetical loadings are derived from a transformation of the First Year Students Data; see Table 12.8. While the DAPPFR transformation was not restricted to being orthogonal, this transformation turned out to being so close to orthogonal that it will be considered here as an orthogonal transformation with the projections on the normals (structure loadings) being taken as the hypothetical loadings for all cases to be considered for oblique factor transformations. The factor transformations to these hypothetical loadings will be applied to the Fourth Year Students principal factors given in Table 12.7.

12.2.1.1 Raw Weights, Fixed Hypothetical Loadings

This appears to be the first, workable solution to the transformation to hypotheized loadings. Mosier (1959) proposed this as an approximate solution to his development for Normals, Fixed Hypothesized Loadungs which is discussed in section 12.2.1.3. Hurley and Cattell (1962) used this procedure in their deveopment: "The Procrustes program: producing direct rotation to test a hypothesized factor structure." They adopted the term "Procrustes Transformation" following a Greek mythology concerning a highwayman who made all his victims fit his bed. This term has become adopted to apply to all the various procedures to obtain a transformation to hypothesized factor loadings.

A separate solution is obtained for each existant column of H. The criterion for this case is:

$$\xi_{fm} = \sum_{i \in I_m} (v_{im} - h_{im})^2 \quad . \tag{12.5}$$

which may be written with equations (12.4):

Table 12.7

Principal Factors of Scaled Covariance Matrices*
Medical Students Data

First Year Students			Fourth Year Students		
Factor Loadings			Factor Loadings		
	1	2		1	2
1	.74	.569	1	.535	.212
2	-1.27	.485	2	.518	.465
3	-.006	.393	3	.552	.449
4	.844	.004	4	.770	-.397
5	.851	.022	5	.749	-.395

*Principal factors of scaled covariance matrices using Joreskog and Goldberger (1972) initial uniqueness.

Table 12.8
Factor Transformation* of First year medical Students Data

Structure Loadings Solution

Normal Vectors

	1	2
1	-.001	1.000
2	1.000	.029

Projections on Normals
(Structure Loadings)

	1	2
1	.568	.090
2	.486	-.113
3	.393	.005
4	-.005	.844
5	.013	.852

Cosines of Angles
Between Normals

	1	2
1	1.000	.018
2	.018	1.000

Pattern Loadings Solution

Trait Vectors

	1	2
1	-.029	1.000
2	1.000	.011

Factor Weights
(Pattern Loadings)

	1	2
1	.568	.091
2	.486	-.113
3	.393	.005
4	-.005	.844
5	.013	.852

Factor
Intercorrelations

	1	2
1	1.000	-.018
2	-.018	1.000

*Dapper transformation(see Chapter 11).

$$\xi_{fm} = \sum_{i \in I_m} \left(\sum_{j=1}^{r} a_{ij} w_{jm} - h_{im} \right)^2 \quad . \tag{12.6}$$

This is an ordinary least squares problem for which the standard solution for the weights vector is:

$$\underline{w}_m = \left(\tilde{A}'_m \tilde{A}_m \right)^{-1} \left(\tilde{A}'_m \underline{\tilde{h}}_m \right) \quad . \tag{12.7}$$

Note that there is an implied restriction that $\left(\tilde{A}'_m \tilde{A}_m \right)$ not be singular so as to possess an inverse. Also, a degenerate solution is obtained when all defined hypothetical loadings for one column of H are zero. In this case the weights vector $\underline{w}_m$ will contain zeros.

This solution is frequently used with an arbitrary additional step of scaling the derived vector of transformation weights, $\underline{w}_m$, either to unit length so as to be a normal to an hyperplane or to being a column of matrix T^{-1}. This rescaling of the vector $\underline{w}_m$ violates the least squares solution of criterion ξ_{fm}. However, this combination of operations may be justified with the canonical type criterion to be described subsequently.

Table 12.9 presents result for the Fourth Year Medical Students Data with the hypothetical loadings being given at the left. The raw weights solution is given in the middle of Table 12.9 with these weights being obtained by equation 12.7. The squared lengths of the raw weights are given in the diagonal of the product matrix. Note that the squared length of the first weights vector is markedly less than unity while the squared length of the second weights vector is larger than unity. The normalized weights solution on the right is obtained by scaling the weights vectors to unit length as shown in the diagonal of the product matrix. The sums of squared differences are given in the bottom line of the table. Note that the normalization procedure has increased these sums of squared differences. An important point is that the solution is no longer orthogonal as shown in the off-diagonal entry in the product matrix. There has been a change in the obliqueness of the structure from the first year students to the fourth year students. This effect will be noted for the orthogonal transformations to follow.

As discussed previously, this criterion is not scale free. Consider that the hypothesized weights were doubled. The raw transformation weights would be doubled as would be the loadings on these raw weights. However, the normalization procedure would compensate for these changes and yield the same solution as before the doubling. While the normalization process is an arbitrary addition to the solution, it does provide a scale free aspect to the revised solution.

12.2.1.2 Raw Weights, Proportional Hypothetical Loadings.

Table 12.9
Solution For: Raw Weights, Fixed Hypothetical Loadings
Fourth Year Medical Students Data

Hypothetical Loadings | Transformatin Results: Raw Weights | Normalized Weights

Transformation Weights Matrices

	Raw 1	Raw 2		Normalized 1	Normalized 2
1	.386	.636	1	.500	.575
2	.669	-.906	2	.866	-.818

Given Loadings / Transformed Loadings

	Given 1	Given 2		Raw 1	Raw 2		Normalized 1	Normalized 2
1	.568	.090	1	.349	.148	1	.451	.134
2	.486	-.113	2	.511	-.092	2	.662	-.083
3	.393	.005	3	.514	-.055	3	.665	-.050
4	-.005	.844	4	.032	.850	4	.041	.767
5	.013	.852	5	.025	.834	5	.033	.754

Product Matrices Of
Transformation Weights Matrices

	Raw 1	Raw 2		Normalized 1	Normalized 2
1	.597	-.360	1	1.000	-.421
2	-.360	1.225	2	-.421	1.000

Sums of Squared Differences

Raw 1	Raw 2	Normalized 1	Normalized 2
.065	.008	.121	.021

A separate solution could be contemplated for each existant column of H . The criterion for this case is:

$$\xi_{pm} = \sum_{i \in I_m} (v_{im} - h_{im} c_m)^2 \qquad (12.8)$$

where c_m is an unkown constant of proportionality to be determined in the solution. The minimum solution is a degenerate case with vector $\underline{w}_m = 0$ and $c_m = 0$. This is an unusable criterion.

12.2.1.3 Normals, Fixed Hypothetical Loadings.

This is the case considered by Mosier (1959) in his article "Determining a simple structure when loadfings for certain tests are known". A separate solution is obtained for each existant column of H . The criterion for this situation is the same as for Raw Weights, Fixed Hypothetical Loaings with a constraint that the weight vector be of unit length. Similar to equation (12.6):

$$\zeta_{fm} = \sum_{i \in I_m} \left(\sum_{j=1}^{r} a_{ij} w_{jm} - h_{im} \right)^2 \qquad (12.9)$$

with the constraint that:

$$\sum_{j=1}^{r} w_{jm}^2 = 1 \quad . \qquad (12.10)$$

A criterion which combines the functions in equations (12.9) and (12.10) utilizes a LaGrange multiplier, β_m , is:

$$\overset{\#}{\zeta}_{fm} = \sum_{i \in I_m} \left(\sum_{j=1}^{r} a_{ij} w_{jm} - h_{im} \right)^2 - \beta_m \left(\sum_{j=1}^{r} w_{jm}^2 - 1 \right) \quad . \qquad (12.11)$$

This equation may be expressed in matrix form as:

$$\begin{aligned} \overset{\#}{\zeta}_{fm} &= \underline{w}_m' \tilde{A}_m' \tilde{A}_m \underline{w}_m - 2 \underline{w}_m \tilde{A}_m' \underline{\tilde{h}}_m + \underline{\tilde{h}}_m' \underline{\tilde{h}}_m - \underline{w}_m' \beta_m \underline{w}_m + \beta_m \quad , \\ &= \underline{w}_m' \left(\tilde{A}_m' \tilde{A}_m - \beta_m I \right) \underline{w}_m - 2 \underline{w}_m' \tilde{A}_m' \underline{\tilde{h}}_m + \underline{\tilde{h}}_m' \underline{\tilde{h}}_m + \beta_m \quad . \end{aligned} \qquad (12.12)$$

To minimize function $\overset{\#}{\zeta}_{fm}$, the partial derivatives are found with respect to the elements, w_{jm} , of the weights vector and are set equal to zero.

$$\frac{\partial \overset{\#}{\zeta}_{fm}}{\partial w_{jm}} = 2 \sum_{i \in I_m} \left(\sum_{k=1}^{r} a_{ik} w_{km} - h_{im} \right) a_{ij} - 2 \beta_m w_{jm} = 0 \quad . \qquad (12.13)$$

This equation may be written in matrix form for all entries in $\underline{w}_m$ as:

$$2 \tilde{A}_m' \tilde{A}_m \underline{w}_m - 2 \tilde{A}_m' \underline{\tilde{h}}_m - 2 \beta_m \underline{w}_m = 0 \quad . \qquad (12.14)$$

With algebraic operations this equation may be written as:

$$\left(\tilde{A}_m' \tilde{A}_m - \beta_m I \right) \underline{w}_m = \overset{\#}{A}{}_m' \underline{\tilde{h}}_m \quad . \qquad (12.15)$$

Conveient definitions follow. Let:

$$X_m = \tilde{A}'_m \tilde{A}_m \quad ; \tag{12.16}$$

$$\underline{y}_m = \tilde{A}'_m \underline{\tilde{h}}_m \quad . \tag{12.17}$$

Then:

$$(X_m - \beta_m I)\,\underline{w}_m = \underline{y}_m \quad . \tag{12.18}$$

The solution for the restricted criterion $\overset{\#}{\zeta}_{fm}$ is:

$$\underline{w}_m = (X_m - \beta_m I)^{-1}\underline{y}_m \quad . \tag{12.19}$$

The Lgrange multiplier, β_m , is to be determined to statisfy the restriction in equation (12.10).

In the solution for β_m , a continuous function q_m is defined by:

$$q_m = \underline{w}'_m \underline{w}_m$$

which with equation (12.19) becomes:

$$q_m = \underline{y}'_m (X_m - \beta_m I)^{-2}\underline{y}_m \quad . \tag{12.20}$$

Thus, q_m is a function of β_m . Solution of this function for the values of β_m for which q_m equals unity as per equation (12.10) is facilitated by a transformation defined from the eigensolution for matrix X_m given in the following equation:

$$X_m = V_m \Lambda_m V'_m \tag{12.21}$$

where Λ_m is a diagonal matrix of eigenvalues and V_m is a square, orthonormal matrix of eigenvectors. Note that:

$$(X_m - \beta_m I) = V_m(\Lambda_m - \beta_m I)V'_m \quad ; \tag{12.22}$$

$$(X_m - \beta_m I)^{-2} = V_m(\Lambda_m - \beta_m I)^{-2}V'_m \quad . \tag{12.23}$$

Equation (12.20) may be written as:

$$q_m = \underline{y}'_m V_m(\Lambda_m - \beta_m I)^{-2}V'_m \underline{y}_m \quad . \tag{12.24}$$

A notational simplification involves defining $\underline{z}_m$ by:

$$\underline{z}_m = V'_m \underline{y}_m \quad . \tag{12.25}$$

Then:

$$q_m = \underline{z}'_m (\Lambda_m - \beta_m I)^{-2}\underline{z}_m \quad . \tag{12.26}$$

With $(\Lambda_m - \beta_m I)$ being a diagonal matrix, equation (12.26) may be converted to sumational notation:

$$q_m = \sum_{j=1}^{r} z_{jm}^2 \Big/ (\lambda_{jm} - \beta_m)^2 \quad . \tag{12.27}$$

Note that q_m is the sum of r terms each of which must be non-negative due to the squares of the numerator and the denominator. Thus, q_m must be non-negative. Further inspection of this equation indicates that q_m will approach zero when β_m approaches either $+\infty$ or $-\infty$. Another point is that term j in the sum will approach $+\infty$ when β_m approaches λ_{jm} for $j = 1$

to r . At these points, q_m approaches $+\infty$. Further properties of this function is provided by the derivative of q_m with respect to β_m :

$$\frac{d\,q_m}{d\,\beta_m} = 2\sum_{j=1}^{r} z_{jm}^2 \Big/ (\lambda_{jm} - \beta_m)^3 \quad . \tag{12.28}$$

When β_m is less than λ_{rm} (λ_{rm} being the least eigenvalue of X_m) this derivative is positive so that the function in the range of β_m from $-\infty$ to λ_{rm} starts equal to zero and increases to ∞ as β_m approaches λ_{rm} . Consequently, there is a solution for q_m equal to unity in this range. There is another solution in the range of β_m from λ_{1m} to ∞ . Also, there are possible solutions in the spaces of β_m between consecutive eigenvalues. Study of these possible solutions leads to rejecting them for the desired minimization of the criterion $\overset{\#}{\zeta}_{fm}$. The desired solution is for β_m in the range from $-\infty$ to λ_{rm} . An iterative solution in this range is described in the following paragraph.

For the iterative procedure, equation (12.27) is written as:

$$q_m = z_{rm}^2 \Big/ (\lambda_{rm} - \beta_m)^2 + \sum_{j=1}^{r-1} z_{jm}^2 \Big/ (\lambda_{jm} - \beta_m)^2 \tag{12.29}$$

in which the final term of the sum is pulled out as a separate term. This equation is somewhat similar to an hyperbolic function which may be written as:

$$y = \frac{a}{(b-x)^2} + c \tag{12.30}$$

in which q_m is replaced by y , z_{rm}^2 is replaced by a, λ_{rm} is replaced by b , and β_m is replaced by x . The final term of equation (12.29) is considered, approximately, to be a constant c . The range of x is considered to be from $-\infty$ to λ_{rm} . Note that as $x \to -\infty$, $y \to c$ so that there is a left hand asymptote of c . The derivative of y with respect to x is:

$$\frac{dy}{dx} = y' = \frac{2a}{(b-x)^3} \quad . \tag{12.31}$$

The plan of the iterative procedure is to start from some trial value of x , termed x_t , then to define the parameters for the function in equation (12.30) for this value of x_t . Given the function of equation (12.30), a solution for x_1 is made for $y_1 = 1$. The obtained x_1 becomes the next trial value of x_t . Due to the similarity of equation (12.30) to the function for q_m in equation (12.29) this procedure should converge to the desired value of β_m . Parameter b is set equal to λ_{rm} . Trial values of parameters a and c are designated by a_t and c_t which are developed from solution of equations (12.30) and (12.31) for trial values y_t and y'_t derived from a trial value of $\beta_m = x_t$. y_t is the value of q_m obtained from equation (12.27) with the trial value of β_m and y'_t , the trial value of the derivative in equation (12.28) for the trial value of β_m . Then:

$$a_t = \tfrac{1}{2}\, y'_t\, (b - x_t)^3 \quad ; \tag{12.32}$$

$$c_t = y_t - \frac{a_t}{(b - x_t)^2} \quad . \tag{12.33}$$

The solution for x_1 when $y_1 = 1$ involves a solution from equation (12.30):

$$x_1 = b - \sqrt{\frac{a_t}{(1 - c_t)}} \quad . \tag{12.34}$$

In the range of $\beta_m = x_t$ from $-\infty$ to λ_{rm} the derivative in equation (12.28) must be positive and the term $(b - x_t)^3$ of equation (12.32) must be positive so that a_t given in equation (12.32) must be positive. It is imperative that $(1 - c_t)$ in equation (12.34) be positive. A computer protection when $(1 - c_t)$ is negative is to set it equal to a small positive value. A further condition in the program is to make sure that the obtained value of y_1 is nearer to 1 than the input trial value, y_t. When this is not so, a smaller change in x_t should be tried.

Table 12.10 presents the illustration of this transformation for the Fourth Year Medical Students Data. These results are similar to those presented in Table 12.9 for the Raw Weights, Fixed Hypothetical Loadings case with the sums of squared differences between those for the raw weights and the normalized weights given in Table 12.9. The squared lengths of the weight vectors given in the product matrix are unity corresponding to the constraint given in equation (12.10).

Mosier (1959) suggested an approximate solution obtained by setting β_m equal to zero and then normalizing the resulting weight vectors. This is identical to the procedure given earlier in section 12.2.1.1 for the Raw Weights, Fixed Hypothetical Loadings case.

For the special case when all defined hypothetical loadings for a factor equal zero the solution degenerates to the least square hyperplane fitting described in Chapter 10. In this case vector $\underline{\tilde{h}}_m$ equals zero as does vector $\underline{y}_m$ as defined in equation (12.17). Then equation (12.18) becomes the eigen problem of equation (10.26) with β_m being the least eigenvalue of X_m and $\underline{w}_m$ being the corresponding eigenvector.

12.2.1.4 Normals, Proportional Hypothetical Loadings.

This is a companion case to the preceding case with, now, the hypothetical loadings being subject to a possible constant of proportionality for each transformed factor. A separate solution is obtained for each existant column of H. The major criterion for this case is:

$$\zeta_{pm} = \sum_{i \in I_m} \left(\sum_{j=1}^{r} a_{ij} w_{jm} - h_{im} c_m \right)^2 \tag{12.35}$$

where c_m is a constant of proportionality. The constraint that vector $\underline{w}_m$ is of unit length is given in equation (12.10). Note the similarity of this criterion with the criterion for the preceding case in equations (12.9) and (12.10). The added feature is the constant of proportionality for the hypothetical loadings. A criterion which combines the functions in equation (12.35) and (12.10) utilizes a LaGrange multiplier, β_m, is:

Table 12.10
Solution For: Normals, Fixed hypothetical Loadings*
Fourth Year Medical Students Data

Hypothetical Loadings

Given Loadings

	1	2
1	.568	.090
2	.486	-.113
3	.393	.005
4	-.005	.844
5	.013	.852

Transformatin Results

Transformation Weights
(Normal Vectors)

	1	2
1	.429	.604
2	.903	-.797

Projections on Normals
(Structure Loadings)

	1	2
1	.421	.154
2	.642	-.057
3	.643	-.024
4	-.028	.782
5	-.035	.767

Cosines of Angles
Between Normals

	1	2
1	1.000	-.460
2	-.460	1.000

Sums of Squared Differences

1	2
.111	.019

*Mosier's (1939) exact solution.

$$\overset{\#}{\zeta}_{pm} = \sum_{i \in I_m} \left(\sum_{j=1}^{r} a_{ij} w_{jm} - h_{im} c_m \right)^2 - \beta_m \left(\sum_{j=1}^{r} w_{jm}^2 - 1 \right) \quad ; \tag{12.36}$$

$$= \sum_{i \in I_m} \left(\sum_{j=1}^{r} a_{ij} w_{jm} \right)^2 - 2 \sum_{i \in I_m} \left(\sum_{j=1}^{r} a_{ij} w_{jm} \right) h_{im} c_m + \sum_{i \in I_m} h_{im}^2 c_m^2 - \beta_m \left(\sum_{j=1}^{r} w_{jm}^2 - 1 \right) \quad . \tag{12.37}$$

This equation may be expressed in matrix form as:

$$\overset{\#}{\zeta}_{pm} = \underline{w}_m' \tilde{A}_m' \tilde{A}_m \underline{w}_m - 2 c_m \underline{\tilde{h}}_m' \tilde{A}_m \underline{w}_m + c_m^2 \underline{\tilde{h}}_m' \underline{\tilde{h}}_m - \beta_m (\underline{w}_m' \underline{w}_m - 1) \tag{12.38}$$

Minimization of $\overset{\#}{\zeta}_{pm}$ is accomplished using the derivative with respect to c_m. This derivative is set equal to zero.

$$\frac{d\overset{\#}{\zeta}_{pm}}{dc_m} = -2 \sum_{i \in I_m} \left(\sum_{i=1}^{r} a_{ij} w_{jm} \right) h_{im} + 2 c_m \sum_{i \in I_m} h_{im}^2 = 0 \quad . \tag{12.39}$$

Then:

$$c_m = \frac{1}{\sum_{i \in I_m} h_{im}^2} \sum_{i \in I_m} \left(\sum_{j=1}^{r} a_{ij} w_{jm} \right) h_{im} \quad . \tag{12.40}$$

In matrix form:

$$c_m = \frac{1}{\underline{\tilde{h}}_m' \underline{\tilde{h}}_m} \underline{\tilde{h}}_m' \tilde{A}_m \underline{w}_m \quad . \tag{12.41}$$

With the nature of equations (12.40) and (12.41) a restriction on the use of this criterion is that the sum of squares of the hypothesized loadings for each factor must be greater than zero. The next step is to obtain the partial derivative of $\overset{\#}{\zeta}_{pm}$ with respect to element w_{jm} of the weight vector. This partial derivative is set equal to zero.

$$\frac{\partial \tilde{\zeta}_{pm}}{\partial w_{jm}} = 2 \sum_{i \in I_m} \left(\sum_{k=1}^{r} a_{ik} w_{km} \right) a_{ij} - 2 \sum_{i \in I_m} a_{ij} h_{im} c_m - 2 \beta_m w_{jm} = 0 \quad . \tag{12.42}$$

Let $\underline{\tilde{a}}_{jm}$ be column vector of the j'th column of $\tilde{A}_m$. Then equation (12.42) becomes, in matrix form:

$$\underline{\tilde{a}}_{jm}' \tilde{A}_m \underline{w}_m - c_m \underline{\tilde{a}}_{jm}' \underline{\tilde{h}}_m - \beta_m w_{jm} = 0 \quad . \tag{12.43}$$

For all values of j:

$$\tilde{A}_m' \tilde{A}_m \underline{w}_m - c_m \tilde{A}_m' \underline{\tilde{h}}_m - \beta_m \underline{w}_m = 0 \quad . \tag{12.44}$$

Substitution from equation (12.41) into equation (12.44) yields:

$$\tilde{A}_m' \tilde{A}_m \underline{w}_m - \frac{1}{\underline{\tilde{h}}_m' \underline{\tilde{h}}_m} \tilde{A}_m' \underline{\tilde{h}}_m \underline{\tilde{h}}_m' \tilde{A}_m \underline{w}_m - \beta_m \underline{w}_m = 0 \quad . \tag{12.45}$$

Or:

$$\left(\tilde{A}_m' \tilde{A}_m - \frac{1}{\underline{\tilde{h}}_m' \underline{\tilde{h}}_m} \tilde{A}_m' \underline{\tilde{h}}_m \underline{\tilde{h}}_m' \tilde{A}_m - \beta_m I \right) \underline{w}_m = 0 \quad . \tag{12.46}$$

A useful definition is:

$$X_m = \tilde{A}'_m\tilde{A}_m - \frac{1}{\underline{\tilde{h}}'_m\underline{\tilde{h}}_m}\tilde{A}'_m\underline{\tilde{h}}_m\underline{\tilde{h}}'_m A_m \quad . \tag{12.47}$$

Equation (12.46) becomes:

$$(X_m - \beta_m I)\,\underline{w}_m = 0 \quad . \tag{12.48}$$

This equation is in the form of an eigen problem with β_m being an eigenvalue of matrix X_m and $\underline{w}_m$ being the corresponding eigenvector. To satisfy the constraint of equation (12.10) $\underline{w}_m$ is to be scaled to unit length. There is, now, the problem of selecting the desired eigenvalue to minimize $\overset{\#}{\zeta}_{pm}$. Substitute the value of c_m from equation (12.41) into equation (12.38) to obtain:

$$\begin{aligned}\overset{\#}{\zeta}_{pm} &= \underline{w}'_m\tilde{A}'_m\tilde{A}_m\underline{w}_m - 2\frac{1}{\underline{\tilde{h}}'_m\underline{\tilde{h}}_m}\underline{w}'_m\tilde{A}'_m\underline{\tilde{h}}_m\underline{\tilde{h}}'_m\tilde{A}_m\underline{w}_m \\ &\quad + \frac{1}{\left(\underline{\tilde{h}}'_m\underline{\tilde{h}}_m\right)^2}\underline{w}'_m\tilde{A}'_m\underline{\tilde{h}}'_m\underline{\tilde{h}}_m\tilde{A}_m\underline{w}_m\underline{\tilde{h}}'_m\underline{\tilde{h}}_m + \beta_m(\underline{w}'_m\underline{w}_m - 1) \\ &= \underline{w}'_m\left(\tilde{A}'_m\tilde{A}_m - \frac{1}{\underline{\tilde{h}}'_m\underline{\tilde{h}}_m}\tilde{A}'_m\underline{\tilde{h}}_m\underline{\tilde{h}}'_m A_m\right)\underline{w}_m + \beta_m(\underline{w}'_m\underline{w}_m - 1) \quad .\end{aligned} \tag{12.49}$$

With the definition of X_m in equation (12.47) and with $\underline{w}_m$ being a unit length vector and the relation between ζ_{pm} and $\overset{\#}{\zeta}_{pm}$ from equations (12.35) and (12.36) equation (12.49) yields:

$$\zeta_{pm} = \overset{\#}{\zeta}_{pm} = \underline{w}'_m X_m \underline{w}_m \quad . \tag{12.50}$$

From the properties of the eigen solution:

$$\zeta_{pm} = \beta_m \quad . \tag{12.51}$$

Then, a minimum value of ζ_{pm} is obtained by using the least eigenvalue of X_m and the corresponding unit length eigenvector as the desired weight vector $\underline{w}_m$. This least eigenvalue is the value of the criterion. Note that the eigenvector may be reversed in sign and still be of unit length. A sugestion is to choose the direction of the eigenvector so that the sum of the transformed loadings will be positive.

Results for the solution for this criterion are given in Table 12.11. Adjustment of the hypothetical loadings is accomplished by the constants of proportionality given at the top of the middle section of the table. Note that the first factor has a constant greater than unity so that the hypothetical loadings are increased while the constant for the second factor is less than unity so that the hypothetical loadings are decreased in absolute value. The transformation results are very similar to results obtained by preceding criteria. As should be expected, the inclusion of the constants of proportionality has reduced the sums of squared differences from those in Table 12.10 for Normals, Fixed Hypothetical Loadings.

12.2.1.5 Trait Vectors, Fixed Hypothetical Loadings.

Table 12.11
Solution For: Noramls, Proportional Hypothetical Loadings
Fourth Year Medical Students Data

Hypothetical Loadings — Transformation Results: Raw Weights / Normalized Weights

Transformation Weights Matrices

Constants of Proportionality

1	2
1.155	.898

Transformation Weights (Normal Vectors)

	1	2
1	.469	.573
2	.883	-.820

Given Loadings

	1	2
1	.568	.090
2	.486	-.113
3	.393	.005
4	-.005	.844
5	.013	.852

Adjusted Loadings

	1	2
1	.656	.081
2	.561	-.101
3	.454	.004
4	-.006	.758
5	.015	.765

Projections on Normals (Structure Loadings)

	1	2
1	.438	.133
2	.654	-.084
3	.655	-.052
4	.010	.767
5	.002	.753

Cosines of Angles Between Normals

	1	2
1	1.000	-.455
2	-.455	1.000

Sums of Squared Differences

1	2
.097	.006

With this function there is a switch from minimizing differences between projections on normals (structure loadings) and the hypothetical loadings to differences between factor weights (pattern loadings) and the hypothetical loadings. For the present case the major criterion is:

$$\Psi_f = \sum_{m=1}^{r} \left\{ \sum_{i \in I_m} (b_{im} - h_{im})^2 \right\} \qquad (12.52)$$

where:

$$B = AT^{-1} \quad ; \qquad (7.4)$$

$$R_{bb} = TT' = \Phi \qquad (12.53)$$

Rows of matrix T contain the trait vectors as described in Chapter 7 with the trnasfomred factor weights (pattern loadings) being in matrix B. All columns of the hypothetical loadings in matrix H must be defined. However, not all entries in each column of H have to be specified. Matrix R_{bb} of correlations among the transformed factors is defined in equation (7.1) and is renamed here for conveience as Φ. Since this is a correlation matrix the diagonal entries must be unity so that:

$$\mathrm{Diag}(\Phi) = I \quad . \qquad (12.54)$$

Criterion Ψ_f is to be minimized under the constraint given in equation (12.54).

Gruvaeus (1970) described a procedure to accomplish the minimization using a series of Fletcher & Powell (1963) minimizing transformations. Browne (1972) described an alternative method to obtain the minimum solution using a series of elemental transformations. In each of these elemental transformations a single trait is considered as a pivot trait while another trait is considered as a spoke trait. The pivot trait is transformed in the direction of the spoke trait. A subcycle consists of transformations of the pivot trait in the direction of each of the other traits as the spoke trait. A cycle of elemental transformations consists of considering each of the traits in turn as the pivot trait. These elemental transformations are continued until every elemental transformation in a cycle is less than some prescribed minimum transformation. Following is a description of a single elemental transformation.

Let the pivot trait be deignated j and the spoke trait be designated k. Then the criterion Ψ_f can be written as:

$$\Psi_f = \sum_{i \in I_j} (b_{ij} - h_{ij})^2 + \sum_{i \in I_k} (b_{ik} - h_{ik})^2 + \sum_{\substack{m=1 \\ m \neq j,k}}^{r} \left\{ \sum_{i \in I_m} (b_{im} - h_{im})^2 \right\} \qquad (12.55)$$

Let $\underline{t}_j$ and $\underline{t}_k$ be trait vectors j and k which are rows j and k of the trait matrix T. Also, let $\tilde{\underline{t}}_j$ be the transformed trait vector j. The transformation is given below

$$s_{jj}\underline{t}_j + s_{jk}\underline{t}_k = \tilde{\underline{t}}_j \qquad (12.56)$$

where s_{jj} and s_{jk} are the transformation coefficients. This transformation is similarr to the transformation in the direct oblimin procedure described in Chapter 11, see equation (11.59). The restriction that $\tilde{\underline{t}}_j$ be a unit vector leads to the following restriction on the transformation coefficients.

$$s_{jj}^2 + 2s_{jj}s_{jk}\phi_{jk} + s_{jk}^2 = 1 \qquad (12.57)$$

where ϕ_{jk} is the correlation between traits j and k in matrix Φ . These coefficients can be recorded in row j and columns j and k of a 'shift' matrix S which has unities in all other diagonal cells and zeros in all other off-diagonal cells. This transformation is illustrated in expanded form for a 4 factor case with $j = 1$ and $k = 2$.

$$\begin{bmatrix} s_{11} & s_{12} & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix} \begin{bmatrix} \underline{t}_1 \\ \underline{t}_2 \\ \underline{t}_3 \\ \underline{t}_4 \end{bmatrix} = \begin{bmatrix} \tilde{\underline{t}}_1 \\ \underline{t}_2 \\ \underline{t}_3 \\ \underline{t}_4 \end{bmatrix}$$

In matrix equation form:

$$ST = \tilde{T} \quad . \qquad (12.58)$$

The transformed factor weights matrix, $\tilde{B}$, is given by:

$$\tilde{B} = A\tilde{T}^{-1}$$

which with equations (7.4) and (12.58) becomes

$$\tilde{B} = AT^{-1}S^{-1} = B\,S^{-1} \quad . \qquad (12.59)$$

Matrices S and S^{-1} are illustrated in expanded form for the 4 factor situation used previously.

$$\begin{bmatrix} s_{11} & s_{12} & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{s_{11}} & \frac{-s_{12}}{s_{11}} & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix} = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix}$$

Columns of matrices B and $\tilde{B}$ are treated as column vectors such as $\underline{b}_j$ and $\tilde{\underline{b}}_j$. Then, equation (12.59) may be illustrated as follows.

$$[\,\underline{b}_1 \quad \underline{b}_2 \quad \underline{b}_3 \quad \underline{b}_4\,] \begin{bmatrix} \frac{1}{s_{11}} & \frac{-s_{12}}{s_{11}} & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix} = [\,\tilde{\underline{b}}_1 \quad \tilde{\underline{b}}_2 \quad \underline{b}_3 \quad \underline{b}_4\,]$$

Note in the illustration that factor weights are transformed for both factor 1 and 2 . In the general case:

$$\underline{b}_j \frac{1}{s_{jj}} = \tilde{\underline{b}}_j \quad ; \qquad (12.60)$$

$$\underline{b}_j \frac{-s_{jk}}{s_{jj}} + \underline{b}_k = \tilde{\underline{b}}_k \quad . \qquad (12.61)$$

The restriction that the transformed trait vector $\tilde{\underline{t}}_j$ be of unit length resulted in the restriction of equation (12.57). A transformation of variables is used to implement this restriction. Define:

$$z_{jk} = \frac{s_{jk}}{s_{jj}} \quad . \qquad (12.62)$$

Solutions for s_{jj} and s_{jk} follow using equation (12.57).

$$\frac{1}{s_{jj}^2} = 1 + 2\phi_{jk}z_{jk} + z_{jk}^2 \quad ; \tag{12.63}$$

$$\frac{1}{s_{jj}} = \sqrt{1 + 2\phi_{jk}z_{jk} + z_{jk}^2} \quad . \tag{12.64}$$

In order to maintain the orientation of trait j the positive square root is used.

$$s_{jk} = s_{jj}z_{jk} \quad . \tag{12.65}$$

Elements of the transformed factor weights follow.

$$\tilde{b}_{ij} = b_{ij}\sqrt{1 + 2\phi_{jk}z_{jk} + z_{jk}^2} \quad . \tag{12.66}$$

$$\tilde{b}_{ik} = b_{ik} - b_{ij}z_{jk} \quad . \tag{12.67}$$

The function to be minimized for this transformation is writen from equation (12.55). (Since the last term is not affected by the transformation of trait j in terms of trait k this term is dropped.)

$$\tilde{\psi}_{fjk} = \sum_{i\in I_j}\left(\tilde{b}_{ij} - h_{ij}\right)^2 + \sum_{i\in I_k}\left(\tilde{b}_{ik} - h_{ik}\right)^2 \quad . \tag{12.68}$$

Expansion of the squared terms with substitution from equations (12.66) and (12.67) with algebraic operations leads to:

$$\tilde{\psi}_{fjk} = a_2 z_{jk}^2 + a_1 z_{jk} + a_* \big/ s_{jj} + a_0 \tag{12.69}$$

where:

$$a_2 = \sum_{i\in I_j} b_{ij}^2 + \sum_{i\in I_k} b_{ij}^2 \quad ; \tag{12.70.2}$$

$$a_1 = 2\left(\sum_{i\in I_j} b_{ij}^2\phi_{jk} - \sum_{i\in I_k} b_{ij}b_{ik} + \sum_{i\in I_k} b_{ij}h_{ik}\right) \quad ; \tag{12.70.1}$$

$$a_* = -2\sum_{i\in I_j} b_{ij}h_{ij} \quad ; \tag{12.70.*}$$

$$a_0 = \sum_{i\in I_j} b_{ij}^2 + \sum_{i\in I_k} b_{ik}^2 - 2\sum_{i\in I_k} b_{ik}h_{ik} + \sum_{i\in I_j} h_{ij}^2 + \sum_{i\in I_k} h_{ik}^2 \quad . \tag{12.70.0}$$

Since coefficient a_2 involves only sums of squares of trial factor weights it, almost always, will be positive. Only in an extremely rare case will a_2 be zero and never negative. Consequently, a_2 will be taken as positive. Coefficient a_1 may be either positive or negative. Coefficient a_* involves the sum of products between the trial factor weights and hypothetical loadings. The sum of squared differences between the trial factor weights and the hypothetical loadings to be small, the sum of products between the trial factor weights and the hypothetical loadings should be positive. If this is not the case, the trial trait vector j should be reversed in sign. This reversal in sign will result in changes in signs of the entries in row j and column j of matrix Φ . With the changes in both the row and the column, the diagonalt entry ϕ_{jj} will remain a positive unity. Also, the entries in column j of T^{-1} will change. Consequently, coefficient a_* will be considered as being negative.

An important special case occurs when a_* equals zero, then equation (12.69) reduces to a parabola with an optimum at:

$$z_{jk} = -a_1 \big/ 2a_2 \quad \text{when } a_* = 0 \quad . \tag{12.71}$$

This case is most likely to occur when all specified hypothetical loadings are zero in which case the present procedure becomes the complement of the least square hyperplane fitting procedure discussed in Chapter 10. The least squares is applied to the factor weights rather than to the projections on the normals. This is a transformation discussed by Lawley and Maxwell (1964) in their article on "Factor Transformation Methods." This solution may be termed "Least Squares Zero Pattern Fitting" (LSQZPF).

Solution of equation (12.69) for z_{jk} to yield a minimum $\tilde{\psi}_{fjk}$ is moderately complex. Except for the case when a_* equals zero there may be two minima so that there is a problem of selecting the most desirable minimum. Browne (1972) suggests using Bailey's (see McCalla, 1967, p. 90) modification of the Newton-Raphson iterative procedure. For this method the first three derivatives of $\tilde{\psi}_{fjk}$ with respect to z_{fjk} are needed. Note: s_{jj} is determined by equation (12.64) using the positive square root. Let:

$$\psi' = \frac{d\tilde{\psi}_{fjk}}{dz_{jk}} = 2a_2 z_{jk} + a_1 + a_*(\phi_{jk} + z_{jk})\, s_{jj} \quad ; \tag{12.72}$$

$$\psi'' = \frac{d\psi'}{dz_{jk}} = 2a_2 + a_*\left(1 - \phi_{jk}^2\right) s_{jj}^3 \quad ; \tag{12.73}$$

$$\psi''' = \frac{d\psi''}{dz_{jk}} = -3a_*\left(1 - \phi_{jk}^2\right)(\phi_{jk} + z_{jk})\, s_{jj}^5 \quad . \tag{12.74}$$

Consider trial t given z_{jkt} . The value of s_{jj} is determined by equation (12.64). Derivatives ψ_t', ψ_t'' and ψ_t''' are determined from equations (12.72), (12.73) and (12.74) using the value of z_{jkt} . The resulting trial z is $z_{jk(t+1)}$ biven by:

$$z_{jk(t+1)} = z_{jkt} - \frac{\psi_t'}{\left\{\psi_t'' - \psi_t'\psi_t''' \big/ 2\psi_t''\right\}} \quad . \tag{12.75}$$

An inspection of the equations yields profitable relations. From equation (12.64). the coefficient s_{jj} approaches 0 as z_{jk} approaches either $-\infty$ or $+\infty$ so that at these extremes the criterion $\tilde{\psi}_{|fjk}$ approaches zero as seen form equation (12.69). However, as indicated in equation (12.68), this criterion must be positive between these two extremes. For a minimum, the second derivative, ψ'' must be positive. A minimum value of this derivative is found by setting the third derivative, ψ''' , equal to zero which occurs when $z_{jk} = -\phi_{jk}$. Thus, the second derivative is positive at all times when the value of ψ'' is positive at this minimum. Then there is only one minimum for the function $\tilde{\psi}_{fjk}$. While there may be no worry that there is a maximum of $\tilde{\psi}_{fjk}$

between two minima, the program should check the algebraic sign of ψ'' at each trial and if it is negative to take a gradient type step toward a minimum.

Solution for a minimum value of criterion Ψ_f starts from some desirable first approximation and makes improvements for each pair of traits as pivot trait and spoke trait until no improvement occurs. A good first approximation may be to obtain trial normals by the simple procedure for raw weight vectors with fixed hypothetical loadings. In case all specified hypothetical loadings for a factor are zero, the normal can be computed by the LSQHYP procedure described in Chapter 10. This trial normals matrix may be converted to a trials trait matrix by the standard formula:

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

where:

$$D = \left[\mathrm{Diag}(FF')^{-1}\right]^{-1/2} \quad . \tag{10.5}$$

Table 12.12 presents an illustrationof the results for the Trait Vectors, Fixed Hypothetical Loadings transformation for the Fourth Year Medical Students Data. Hypothetical loadings given on the left are from the DAPPFR solution for the First Year Medical Students Data given in Table 12.8. While the correlation between factors 1 and 2 was almost zero for the First Year Medical Students, this correlation given in Table 12.12 is .360 for the Fourth Year Medical Students. Thus, while the factors were almost orthogonal for the First Year Medical Students, the factors are fairly correlated for the Fourth Year Medical Students. This is a result to be considered for the orthogonal transformation procedures.

12.2.1.6 Trait Vectors, Proportional Hypothetical Loadings.

This function is very similar to the previous function with the only change being that the hypothetical loadings are multiplied by a constant of proportionality for each factor.

$$\Psi_p = \sum_{m=1}^{r} \left\{ \sum_{i \in I_m} (b_{im} - c_m h_{im})^2 \right\} \tag{12.76}$$

where c_m is a constant of proportionality for each factor applied to the hypothetical loadings for that factor. The discussion following equation (12.52) applies to the present function with the b_{im} being entries in the matrix B and there being a restriction on the trait vectors given in equation (12.54). Again, each trait vector in turn will be considered a s a pivot trait. For each pivot trait, each of the other trait vectors will be considered in turn as a spoke trait. For each pair of a pivot trait and a spoke trait a transformation will be accomplished to a minimum of function Ψ_p . Computations will cycle through the combinations of pivot trait and spioke trait until only

Table 12.12
Solution For: Trait Vectors, Fixed Hypothetical Loadings
Fourth Year Medical Students Data

Hypothetical Loadings

Given Loadings

	1	2
1	.568	.090
2	.486	-.113
3	.393	.005
4	-.005	.844
5	.013	.852

Transformation Results

Trait Vectors

	1	2
1	.469	.573
2	.883	-.820

Factor Loadings
(Pattern Loadings)

	1	2
1	.475	.196
2	.704	-.023
3	.707	.013
4	.024	.857
5	.015	.841

Factor
Intercorrelations

	1	2
1	1.000	.360
2	.360	1.000

Sums of Squared
Differences

	1	2
	.155	.020

Criterion ψ_f = .175

minimal changes occur. Convergence is guaranteed since the function is positive, or at most zero, in nature.

Let the pivot trait be designated j and the spoke trait be designated k. Then the criterion Ψ_p can be written as:

$$\Psi_p = \sum_{i \in I_j} (b_{ij} - c_j h_{ij})^2 + \sum_{i \in I_k} (b_{ik} - c_k h_{ik})^2 + \sum_{\substack{m=1 \\ m \neq j,k}}^{r} \left\{ \sum_{i \in I_m} (b_{im} - c_m h_{im})^2 \right\} \tag{12.77}$$

The transformation of trait vector $\underline{t}_j$ with respect to trait vector $\underline{t}_k$ to the transformed trait vector $\tilde{\underline{t}}_j$ involving transformation coefficients s_{jj} and s_{jk} is given in equation (12.56). The constraint on these coefficients is such that the transformed trait vector $\tilde{\underline{t}}_j$ is of unit length is given in equation (12.57). Transformation of the factor weights $\underline{b}_j$ and $\underline{b}_k$ to transformed factor weights $\tilde{\underline{b}}_j$ and $\tilde{\underline{b}}_k$ is given in equations (12.60) and (12.61). A derived coefficient z_{jk} is defined in equation (12.62) with the relations of s_{jj} and s_{jk} to z_{jk} being given in equations (12.63), (12.64), and (12.65). The transformed factor weights are given in terms of z_{jk} are given in equations (12.66) and (12.67). As in the preceding solution, the function to be minimized for the present transformation is written from equation (12.77). (Since the last term of equation (12.77) is not affected by the transformation of trait j in terms of trait k, this term is dropped.)

$$\tilde{\psi}_{pjk} = \sum_{i \in I_j} (\tilde{b}_{ij} - c_j h_{ij})^2 + \sum_{i \in I_k} (\tilde{b}_{ik} - c_k k_{ik})^2 \quad . \tag{12.78}$$

Substitutions from equations (12.66) and (12.67) with algebraic operations leads to:

$$\begin{aligned} \tilde{\psi}_{pjk} = {} & \sum_{i \in I_j} b_{ij}^2 + 2\phi_{jk} z_{jk} \sum_{i \in I_j} b_{ij}^2 + z_{jk}^2 \sum_{i \in I_j} b_{ij}^2 - 2\, c_j \sqrt{1 + 2\phi_{jk} z_{jk} + z_{jk}^2} \sum_{i \in I_j} b_{ij} h_{ij} + c_j^2 \sum_{i \in I_j} h_{ij}^2 \\ & + \sum_{i \in I_k} b_{ik}^2 - 2 z_{jk} \sum_{i \in I_k} b_{ij} b_{ik} + z_{jk}^2 \sum_{i \in I_k} b_{ij}^2 - 2 c_k \sum_{i \in I_k} b_{ik} h_{ik} + 2 c_k z_{jk} \sum_{i \in I_k} b_{ij} h_{ik} + c_k^2 \sum_{i \in I_k} h_{ik}^2 \quad . \end{aligned} \tag{12.79}$$

A solution for the constants of proportionality in terms of z_{jk} is obtained by setting the derivatives of $\tilde{\psi}_{pjk}$ with respect to these constants of proportionality equal to zero.

$$\frac{d\tilde{\psi}_{pjk}}{dc_j} = -2\sqrt{1 + 2\phi_{jk} z_{jk} + z_{jk}^2} \sum_{i \in I_j} b_{ij} h_{ij} + 2 c_j \sum_{i \in I_j} h_{ij}^2 = 0 \quad . \tag{12.80}$$

$$\frac{d\tilde{\psi}_{pjk}}{dc_k} = 2 \left(z_{jk} \sum_{i \in I_k} b_{ij} h_{ik} - \sum_{i \in I_k} b_{ik} h_{ik} \right) + 2 c_k \sum_{i \in I_k} h_{ik}^2 = 0 \quad . \tag{12.81}$$

These equations lead to:

$$c_j = \sqrt{1 + 2\phi_{jk} z_{jk} + z_{jk}^2} \sum_{i \in I_j} b_{ij} h_{ij} \Big/ \sum_{i \in I_j} h_{ij}^2 \quad ; \tag{12.82}$$

$$c_k = \left(\sum_{i \in I_k} b_{ik} h_{ik} - z_{jk} \sum_{i \in I_k} b_{ij} h_{ik} \right) \Big/ \sum_{i \in I_k} h_{ik}^2 \quad . \tag{12.83}$$

Substitution of these results into equation (12.79) with algebraic operations leads to;

$$\tilde{\psi}_{pjk} = a_2 z_{jk}^2 + a_1 z_{jk} + a_0 \tag{12.84}$$

where:

$$a_2 = \sum_{i\in I_j} b_{ij}^2 + \sum_{i\in I_k} b_{ij}^2 - \frac{\left(\sum_{i\in I_j} b_{ij}h_{ij}\right)^2}{\sum_{i\in I_j} h_{ij}^2} - \frac{\left(\sum_{i\in I_k} b_{ij}h_{ik}\right)^2}{\sum_{i\in I_k} h_{ik}^2} \quad ; \tag{12.85.2}$$

$$a_1 = 2\left\{ \phi_{jk}\sum_{i\in I_j} b_{ij}^2 - \sum_{i\in I_k} b_{ij}b_{ik} - \phi_{jk}\frac{\left(\sum_{i\in I_j} b_{ij}h_{ij}\right)^2}{\sum_{i\in I_j} h_{ij}^2} + \frac{\left(\sum_{i\in I_k} b_{ik}h_{ik}\right)\left(\sum_{i\in I_k} b_{ij}h_{ik}\right)}{\sum_{i\in I_k} h_{ik}^2} \right\} ; \tag{12.85.1}$$

$$a_0 = \sum_{i\in I_j} b_{ij}^2 + \sum_{i\in I_k} b_{ik}^2 - \frac{\left(\sum_{i\in I_j} b_{ij}h_{ij}\right)^2}{\sum_{i\in I_j} h_{ij}^2} - \frac{\left(\sum_{i\in I_k} b_{ik}h_{ik}\right)^2}{\sum_{i\in I_k} h_{ik}^2} \quad . \tag{12.85.0}$$

With equation (12.84) being that of a parabola the minimum solution for $\tilde{\psi}_{pjk}$ occurs at:

$$z_{jk} = -a_1 \big/ 2a_2 \quad . \tag{12.86}$$

Note that the sum of squared hypothetical loadings for each factor must be greater than zero; that is:

$$\sum_{i\in I_m} > 0 \quad \text{for } m = 1, r \,. \tag{12.87}$$

An illustration of the results for the trait vectors, proportional hypothetical loadings transformation is given in Table 12.13. The given hypothetical loadings are the same as used in Table 12.12 from transformation for First Year Medical Students in Table 12.8.. The constants of proportionality obtained in the solution are given at the top of the center section of the Table 12.13 with the adjusted hypothetical loadings being given in the middle of this center section. Note that the hypothetical loadings for the first factor are increased by the constant of proportionality of 1.278 while the hypothetical loadings for the second factor remain practically unchanged. The transformation results are quite similar to those in Table 12.12 for Trait Vectors, Fixed Hypothetical Loadings with the two factors having a correlation of .395. Use of the constants of proportionality has reduced materially the sums of squared differences from Table 12.12 to Table 12.13. Use of the constants of proportionality appears to have a very desirable effect.

12.2.1.7 Canonical Factor Matching.

Table 12.13
Solution For: Trait Vector, Proportional Hypothetical Loadings
Fourth Year Medical Students Data

Hypothetical Loadings | Transformation Results

Constants of Proportionality

1	2
1.278	.992

Trait Vectors

	1	2
1	.799	.867
2	.601	-.498

Given Loadings

	1	2
1	.568	.090
2	.486	-.113
3	.393	.005
4	-.005	.844
5	.013	.852

Adjusted Loadings

	1	2
1	.726	.089
2	.621	-.112
3	.502	.005
4	-.006	.838
5	.017	.846

Factor Weights (Pattern Loadings)

	1	2
1	.490	.165
2	.719	-.066
3	.723	-.029
4	.043	.849
5	.033	.833

Factor Intercorrelations

	1	2
1	1.000	.394
2	.394	1.000

Sums of Squared Differences

1	2
.117	.009

Criterion $\psi_p = .126$

The canonical factor matching function is a normed least squares function with no restraints on the transformation weights nor the constant of proportionality applied to the hypothetical loadings. All six of the preceding matching functions can be classified as raw least squares functions with various constraints on the transformation weights and constants of proportionality. The canonical function is scale free so that the results may be scaled to satisfy various demands. The transformation weights may be scaled to unit vectors so as to be normals to hyperplanes or theses weights may be scaled to be columns of matrix T^{-1} to yield columns of factor weight matrix B. Equation (12.88) gives the canonical factor matching function.

$$\theta_m = \frac{\sum_{i \in I_m} \left(\sum_{j=1}^{r} a_{ij} w_{jm} - h_{im} c_m \right)^2}{\sum_{i \in I_m} \left(\sum_{j=1}^{r} a_{ij} w_{jm} + h_{im} c_m \right)^2} \tag{12.88}$$

The numerator is the same as the raw weights, proportional hypothetical loadings criterion which was an unacceptable criterion; however, the norming by the denominator yields a very acceptable criterion with desirable results. This function follows the development by Tucker (1951) in his report on the synthesis of factor analytic studies in which he used this function in a more complex form as an index of congruence. As will be seen subsequently, this index of congruence is closely related to the coefficient of congruence which Tucker introduced in his report. A solution is made for each factor individually so that an overall coefficient could be considered by summing over the individual values of θ_m when there is a column of hypothetical loadings for every transformed factor.

The matter of scale freeness is considered first using a scaling constant s_m which is any undefined, nonzero constant. Let:

$$\overset{\#}{w}_{jn} = w_{jm} s_m \quad ; \tag{12.89}$$

$$\overset{\#}{c}_m = c_m s_m \quad . \tag{12.90}$$

Also:

$$\overset{\#}{\theta}_m = \frac{\sum_{i \in I_m} \left(\sum_{j=1}^{r} a_{ij} \overset{\#}{w}_{jm} - h_{im} \overset{\#}{c}_m \right)^2}{\sum_{i \in I_m} \left(\sum_{j=1}^{r} a_{ij} \overset{\#}{w}_{jm} + h_{im} \overset{\#}{c}_m \right)^2} \quad . \tag{12.91}$$

With substitution from equations (12.89) and (12.90) a result is obtained that:

$$\overset{\#}{\theta}_m = \theta_m \quad . \tag{12.92}$$

Thus, the function is independent of the scaling defined in equation (12.89) and (12.90).

With the scale free property of the canonical function, this function is not completely identified. A convenient possibility is to define the original constant c_m equal to unity so that the function becomes:

$$\theta_m = \frac{\sum_{i \in I_m} \left(\sum_{j=1}^{r} a_{ij} w_{jm} - h_{im} \right)^2}{\sum_{i \in I_m} \left(\sum_{j=1}^{r} a_{ij} w_{jm} + h_{im} \right)^2} \tag{12.93}$$

The numerator now is the same as the raw weights, fixed hypothetical case. To obtain a minimum solution the partial derivatives of θ_m with respest to the w_{jm}'s are set equal to zero.

$$\frac{\partial \theta_m}{\partial w_{jm}} = \frac{\left\{ \begin{array}{l} \left[\sum_{i \in I_m} \left(\sum_{k=1}^{r} a_{ik} w_{km} + h_{im} \right)^2 \right] \left[2 \sum_{i \in I_m} \left(\sum_{k=1}^{r} a_{ik} w_{km} - h_{im} \right) a_{ij} \right] \\ - \left[\sum_{i \in I_m} \left(\sum_{k=1}^{r} a_{ik} w_{km} - h_{im} \right)^2 \right] \left[2 \sum_{i \in I_m} \left(\sum_{k=1}^{r} a_{ik} w_{km} + h_{im} \right) a_{ij} \right] \end{array} \right\}}{\left[\sum_{i \in I_m} \left(\sum_{k=1}^{r} a_{ik} w_{km} + h_{im} \right)^2 \right]^2} \tag{12.94}$$

Equating these derivatives to zero accompanied with algebraic operations yields:

$$\left\{ \begin{array}{l} \left[\sum_{i \in I_m} \left(\sum_{k=1}^{r} a_{ik} w_{km} \right)^2 + \sum_{i \in I_m} h_{im}^2 \right] \left[\sum_{i \in I_m} a_{ij} h_{im} \right] \\ -2 \left[\sum_{i \in I_m} h_{im} \sum_{k=1}^{r} a_{ik} w_{km} \right] \left[\sum_{i \in I_m} a_{ij} \sum_{k=1}^{r} a_{ik} w_{km} \right] \end{array} \right\} = 0 \quad . \tag{12.95}$$

Then:

$$\sum_{i \in I_m} a_{ij} \sum_{k=1}^{r} a_{ik} w_{km} = \frac{\left[\sum_{i \in I_m} \left(\sum_{k=1}^{r} a_{ik} w_{km} \right)^2 + \sum_{i \in I_m} h_{im}^2 \right]}{2 \left[\sum_{i \in I_m} h_{im} \sum_{k=1}^{r} a_{ik} w_{km} \right]} \sum_{i \in I_m} a_{ij} h_{im} \quad . \tag{12.96}$$

Since the fraction on the left of equation (12.96) is a ratio of scalars, this fraction can be replaced by a coefficient defined in terms of:

$$g_m = \frac{2 \left[\sum_{i \in I_m} h_{im} \sum_{k=1}^{r} a_{ik} w_{km} \right]}{\left[\sum_{i \in I_m} \left(\sum_{k=1}^{r} a_{ik} w_{km} \right)^2 + \sum_{i \in I_m} h_{im}^2 \right]} \quad . \tag{12.97}$$

Then:

$$\sum_{i \in I_m} a_{ij} \sum_{k=1}^{r} a_{ik} w_{km} = \frac{1}{g_m} \sum_{i \in I_m} a_{ij} h_{im} \tag{12.98}$$

which may be written in matrix form as:

$$\tilde{A}_m' \tilde{A}_m \underline{w}_m = \frac{1}{g_m} \tilde{A}_m' \underline{\tilde{h}}_m \quad . \tag{12.99}$$

The solution for $\underline{w}_m$ is given by:

$$\underline{w}_m = \frac{1}{g_m} \left(\tilde{A}_m' \tilde{A}_m \right)^{-1} \left(\tilde{A}_m' \underline{\tilde{h}}_m \right) \quad . \tag{12.100}$$

Note that these transformation weights have the raw weights for the raw weights fixed hypothetical loadings case divided by the coefficient g_m . When the weights are scaled such as to unit length this matter of proportionality is of no consequence. Thus, the present approach justifies the scaling process which was considered to be an arbitrary addition for the raw weights, fixed hypothetical loadings case. The common practice has been justified. Further, since the factor weights in matrix B are proportional by columns to the projections on the normals, the scaling of the weight vectors, $\underline{w}_m$, to columns in matrix T^{-1} is proper.

There are several very interesting relations to be discussed. Following are several useful relations involving conversion to matrix notation and substitution for the value of $\underline{w}_m$ from equation (12.101).

$$\sum_{i\in I_m} h_{im}\sum_{k=1}^{r} a_{ik}w_{km} = \underline{\tilde{h}}'_m\tilde{A}_m\underline{w}_m = \frac{1}{g_m}\left(\underline{\tilde{h}}'_m\tilde{A}_m\right)\left(\tilde{A}'_m\tilde{A}_m\right)^{-1}\left(\tilde{A}'_m\underline{\tilde{h}}_m\right) \quad . \tag{12.101}$$

$$\sum_{i\in I_m}\left(\sum_{k=1}^{r} a_{ik}w_{km}\right)^2 = \sum_{i\in I_m}\left(\sum_{k=1}^{r} w_{km}\right)\left(\sum_{l=1}^{r} a_{il}w_{lm}\right) = \underline{w}'_m\tilde{A}'_m\tilde{A}_m\underline{w}_m \quad ;$$

$$= \frac{1}{g_m^2}\left(\underline{\tilde{h}}'_m\tilde{A}_m\right)\left(\tilde{A}'_m\tilde{A}_m\right)^{-1}\left(\tilde{A}'_m\underline{\tilde{h}}_m\right) \quad . \tag{12.102}$$

$$\sum_{i\in I_m} h_{im}\sum_{k=1}^{r} a_{ik}w_{im} = g_m\sum_{i\in I_m}\left(\sum_{k=1}^{r} a_{ik}w_{im}\right)^2 \quad . \tag{12.103}$$

An important relation is derived by substituting from equation (12.103) into equation (12.97).

$$g_m = \frac{2g_m\left[\sum_{i\in I_m}\left(\sum_{k=1}^{r} a_{ik}w_{km}\right)^2\right]}{\left[\sum_{i\in I_m}\left(\sum_{k=1}^{r} a_{ik}w_{km}\right)^2 + \sum_{i\in I_m} h_{im}^2\right]} \quad .$$

$$\sum_{i\in I_m}\left(\sum_{k=1}^{r} a_{ik}w_{km}\right)^2 + \sum_{i\in I_m} h_{im}^2 = 2\sum_{i\in I_m}\left(\sum_{k=1}^{r} a_{ik}w_{km}\right)^2 \quad .$$

$$\sum_{i\in I_m} h_{im}^2 = \sum_{i\in I_m}\left(\sum_{k=1}^{r} a_{ik}w_{km}\right)^2 \quad . \tag{12.104}$$

This relation states that the sum of squared hypothetical loadings equals the sum of squared raw loadings. This relation is scale free when the scaling relations in equations (12.89) and (12.90) are considered. Remeber that c_m was set equal to zero. With the application of the scaling operations:

$$\sum_{i\in I_m}\left(\overset{\#}{c}_m h_{im}\right)^2 = \sum_{i\in I_m}\left(\sum_{k=1}^{r} a_{ik}\overset{\#}{w}_{km}\right)^2 \quad . \tag{12.105}$$

This equation verifies the scale free nature of the solution. With the result in equation (12.104) the coefficient g_m as given in equation (12.97) may be written as:

$$g_m = \frac{\sum_{i\in I_m} h_{im}\sum_{k=1}^{r} a_{ik}w_{km}}{\sum_{i\in I_m}\left(\sum_{k=1}^{r} a_{ik}w_{km}\right)^2} = \frac{\sum_{i\in I_m} h_{im}\sum_{k=1}^{r} a_{ik}w_{km}}{\sum_{i\in I_m} h_{im}^2} \quad . \tag{12.106}$$

An important alternative is:

$$g_m = \frac{\sum_{i\in I_m} h_{im}\sum_{k=1}^{r} a_{ik}w_{km}}{\sqrt{\sum_{i\in I_m}\left(\sum_{k=1}^{r} a_{ik}w_{km}\right)^2}\ \sqrt{\sum_{i\in I_m} h_{im}^2}} \quad . \tag{12.107}$$

With equation (12.4.V):

$$g_m = \frac{\sum_{i\in I_m} h_{im}v_{im}}{\sqrt{\sum_{i\in I_m} v_{im}^2}\ \sqrt{\sum_{i\in I_m} h_{im}^2}} \quad . \tag{12.108}$$

This is the coefficient of congruence, defined by Tucker (1951), between the transformed factor loadings and the hypothetical loadings. Burt (1948) termed this coefficient as the "unadjusted correlation" between two sets of factor coefficients.

There is an important relation between the index of congruence, θ_m , and the coefficient of congruence, g_m . Expanding equation (12.93):

$$\theta_m = \frac{\sum_{i\in I_m}\left(\sum_{j=1}^{r} a_{ij}w_{jm}\right)^2 - 2\sum_{i\in I_m} h_{im}\sum_{j=1}^{r} a_{ij}w_{jm} + \sum_{i\in I_m} h_{im}^2}{\sum_{i\in I_m}\left(\sum_{j=1}^{r} a_{ij}w_{jm}\right)^2 + 2\sum_{i\in I_m} h_{im}\sum_{j=1}^{r} a_{ij}w_{jm} + \sum_{i\in I_m} h_{im}^2} \quad ;$$

which with equation (12.106) and algebraic operations becomes:

$$\theta_m = \frac{1 - g_m}{1 + g_m} \quad . \tag{12.109}$$

Thus, θ_m and g_m are obversely related in that as g_m increases θ_m decreases so that a maximum value of g_m corresponds to a minimum value of θ_m and vis-a-vers. Korth and Tucker (1976) presented solution for "Procrustes Matching by Congruence Coefficients' in which the coefficient of congrunce was maximized. This is a completely identical solution to the present one.

This approach is termed "Canonical factor matching" due to the considerable analogy to regular canonical analysis in which two batteries of measured attributes are related by weighted composites of the two batteries. The analogy for the present factor matching procedure is to the special case when the second battery is constituted by a single measured attribute. In regular canonical analysis a sample of entities are measured on the attributes in the two batteries. For factor matching the entities are the attributes in the battery being analysed while the measured attributes are the original factors. A major difference is that in regular canonical analysis the measures are adjusted for the mean measures over the sample of entities while in factor matching there is no adjustment for mean values. This shows up in the coefficient of congruence. In the index of congruence given in equation (12.88) the weights for the first battery, the original factor matrix, are the transformation weights while the weight for the second battery, the hypothetical loadings, is the constant of proportionality. For the special case where there is only one measured attribute in the second battery the canonical weights for the first battery are proportional to the multiple regression weights. In the present case the transformation weights are proportional to the least squares solution weights, as given for the Raw weights, fixed hypothetical loadings case.

An aid in the interpretation of the strength of matching is given by a transformation of the coefficient of congruence to an F statistic for random hypothetical loadings. This development involves a series of transformations. Subscripts and special designations are droppeed for the sake of simplicity. Matrix A is taken to involve only rows for which hypothetical loadings are specified. Also, vector h includes only specified entries. From equations (12.102) and (12.104);

$$g^2 = \frac{(h'A)(A'A)^{-1}(A'h)}{h'h} \quad . \tag{12.110}$$

A singular value representation of matrix A follows:

$$A = [P_1 \quad P_2] \begin{bmatrix} \Delta_1 \\ 0 \end{bmatrix} Q' = P\Delta Q' = P_1\Delta_1 Q' \quad . \tag{12.111}$$

Note that P is an n x n orthonormal matrix with sections P_1 , n x r, and P_2 , n x (n−r) , with Δ_1 being an r x r diagonal matrix. Matrix Q is an r x r orthonormal matrix. Usual operations have:

$$A'A = Q\Delta_1^2 Q' \tag{12.112}$$

which is eigen solutione for matrix $A'A$. Matrix P_1 is obtained by:

$$AQ\Delta_1^{-1} = P_1 \tag{12.113}$$

with matrix P_2 being set such that:

$$P_1'P_2 = 0 \quad ; \tag{12.114.1}$$

$$P_2'P_2 = I \quad . \tag{12.114.2}$$

With the definition of P_1 in equation (12.113) it can be shown that:

$$P_1'P_1 = I \tag{12.115}$$

so that matrix P is orthonormal as indicated following equation (12.111). With the eigen solution of equation (12.112):

$$(A'A)^{-1} = Q\Delta_1^{-2}Q'$$

so that, with algebraic operations:

$$A(A'A)^{-1}A' = P_1P_1' \quad . \tag{12.116}$$

The hypothetical loadings vector may be transformed to a vector $\underline{x}$ by:

$$\underline{x} = P'\underline{h}$$

or:

$$\begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \end{bmatrix} = \begin{bmatrix} P_1' \\ P_2' \end{bmatrix} \begin{bmatrix} \underline{h} \end{bmatrix} \quad . \tag{12.117}$$

With this definition and equation (12.116):

$$\underline{h}'A(A'A)^{-1}A'\underline{h} = \underline{x}_1'\underline{x}_1 \tag{12.118}$$

and

$$\underline{h}'\underline{h} = \underline{x}'\underline{x} = \underline{x}_1'\underline{x}_1 + \underline{x}_2'\underline{x}_2 \quad . \tag{12.119}$$

For the present purposes, vector $\underline{h}$ may be considered to be a random vector from a multidimensional normal density function with mean vector equal to 0 and covariance matrix equal to an identity matrix. With this definition, vector $\underline{h}$ is a vector in a random direction in the n dimensional space. The transformation of equation (12.117) expresses this vector in terms of another set of axes. Use of the multidimensional density function the individual entries in vector $\underline{h}$ are independently distributed random normal deviates. After the transformation of equation (12.117) the same is true of the entries in vector $\underline{x}$ so that sums of squares of these entries are distributed as chi squares with degrees of freedom equal to the number of entries entered into the sums. Let:

$$S_1 = \underline{x}_1'\underline{x} = \sum_{i=1}^{r} x_i^2 \tag{12.120.1}$$

$$S_2 = \underline{x}_2'\underline{x}_2 = \sum_{i=r+1}^{n} x_i^2 \tag{12.120.2}$$

From equation (12.110) and equations (12.118) throuhg (12.120.2):

$$g^2 = \frac{S_1}{S_1 + S_2} \quad . \tag{12.121}$$

Algebraic operations yield:

$$\frac{S_1}{S_2} = \frac{g^2}{1 - g^2} \tag{12.122}$$

As seen in equations (12.120.1) and (12.120.2) S_1 and S_2 are sums of independent normal deviates and are, thus, chi squares with r and $(n-r)$ degrees of freedom. From this an F statistic may be defined as:

$$F_{r,n-r} = \frac{(n-r)}{r} \frac{g^2}{1 - g^2} \quad . \tag{12.123}$$

This implies a distribution generated by random vectors $\underline{x}$ and may be used to compare observed coefficients of congruence with such random coefficients. Note that the n is the number of specified hypothetical loadings for a transformed factor.

The preceding F statistic depends on the assumption of the $\underline{h}$ vector being drawn from a multidimensional normal density function. This assumption may not be appropriate and some other multidimensional density function could be used. However, the F statistic may not be accurate, even as an approximation. An alternative for some defined statistical function is the use of Monte Carlo type procedures. One possibility is to run a number of replications using randomly reordered observed $\underline{h}$ vectors. A distribution of coefficients of congruence would be obtained for comparison with the observed coefficient.

Results for an illustrative solution for the canonical transformation system are given in Tables 12.14 and 12.15. In Table 12.14 the hypothetical loadings are the same as those used previously. On the rightof Table 12.14 are the raw transformation results. Of interest are the indices of congruence given toward the bottom of Table 12.14. These appear to be quite small

Table 12.14
Canonical Factor Matching: Raw Results
Fourth Year Medical Students Data

Hypothetical Loadings — Normal Vectors

Transformation Results — Trait Vectors

	1	2
1	.405	.638
2	.702	-.908

Given Loadings

	1	2
1	.568	.090
2	.486	-.113
3	.393	.005
4	-.005	.844
5	.013	.852

Raw Loadings

	1	2
1	.366	.149
2	.536	-.092
3	.539	-.056
4	.034	.852
5	.026	.837

Coefficients Obtained

	1	2
Index of Congruence	.024	.001
Coefficient of Congruence	.953	.997
$F_{2, 3}$	14.8	249
p	.027	.0005

Table 12.15
Canonical Factor Matching: Scaled Results
Fourth Year Medical Students Data

Hypothetical Loadings

Normal Vectors

	1	2
1	.500	.575
2	.866	-.818

Projections on Normals
(Structure Loadings)

	1	2
1	.451	.134
2	.662	-.083
3	.665	-.050
4	.041	.767
5	.033	.754

Cosines of Angles
Between Normls

	1	2
1	1.000	-.421
2	-.421	1.000

Transformation Results

Trait Vectors

	1	2
1	.818	.866
2	.575	-.500

Factor Loadings
(Pattern Loadings)

	1	2
1	.497	.148
2	.730	-.091
3	.733	-.055
4	.046	.846
5	.036	.831

Factor
Intercorrelations

	1	2
1	1.000	.421
2	.421	1.000

which could indicate a good fit of the obtained raw loadings to the hypothetical loadings. Also of interest are the coefficients of congruence which look quite high. However, these coefficients might be compared with the results from the F distribution of equation (12.123). The values of F along with the corresponding p's (proportions above the value of F in the F distribution) are given at the bottom of the table. The F for factor 1 is significant at the 5% level while the F for the second factor is significant well beyond .1% level. The match for factor 1 is not very strong while the match for factor 2 is extremely strong. An interesting alternative is to solve equation (12.123) for the value of the coefficient of congruence for a given level of significance. The value of $F_{2,3}$ for the 1% level of confidence is 30.82 which converts to a g of .977 . This appears to be a quite high value and indicates that the factor matching is very biased to high values. Randon results appear much too good. There should be considerable care taken in interpreting the outcome of factor matching operations. One might conjecture that this observation might apply to all of the methods of factor transformations to hypothetical loadings.

As indicated in preceding discussions the canonical transformation results may be scaled so that the raw weights become normals to hyperplanes or become columns in matrix T^{-1}. The loadings become either projections on the normals (structure loadings) or factor weights (pattern loadings). These scalings involve only constants of proportionality for the various columns of loadings and do not change the coefficients of congruence. The canonical transformation procedure is scale free. Table 12.15 presents the results for the Fourth Year Medical Students data. These are legitimate results due to the scale free property of the canonical transformation approach. Note thatthe structure solution given on the left is identical to the normalized results for the Raw Weights, Fixed Hypothetical Loadings solution given in Table 12.9. Thus, this approach makes legitimate the frequently used Raw Weights, Fixed Hypothetical Loadings procedure.

12.2.2 Orthogonal Factor Transformations

A number of analysts prefer orthogonal transformations to uncorrelated factors due to the mathematical simplicity of the resulting system. Further some analysts argue for statistically independent factors. With uncorrelated factors there is no distinction between normals and trait vectors since these are identical. Also, the structure loadings and pattern loadings are identical. In this case, the matrix Φ is an identity matrix so that:

$$\Phi = TT' = I \quad . \tag{12.124}$$

Then:

$$T' = T^{-1} \tag{12.125}$$

and

$$B = AT^{-1} = AT' \quad . \tag{12.126}$$

In summation form:

$$\phi_{km} = \sum_{j=1}^{r} t_{kj} t_{mj} \tag{12.127}$$

and

$$b_{im} = \sum_{j=1}^{r} a_{ij} t_{mj} \quad . \tag{12.128}$$

Two cases are to be considered: first when the hypothetical loadings are considered to be fixed as given by the analyst and second when each column of hypothetical loadings may be multiplied by a constant of proportionality.

12.2.2.1 Orthogonal Trait Vectors, Fixed Hypothetical Loadings

These transformations started with Green's (1952) development of orthogonal approximations of oblique structures in factor analysis which was followed by Cliff (1966) and Schönemann (1966) who applied Green's results to the approximation of given hypothetical loadings. These procedures applied to fully specified hypothetical loadings matrices. Browne (1972) developed a procedure for orthogonal rotation to a partially specified target (hypothetical loadings matrix). Schönemann's development has several interesting features and will be discussed first to be followed by the discussion of Browne's development.

For the case for a fully specified hypothetical loadings matrix, H, nxr, has all entries in each column specified; thus the, summations over entities in column j may be expressed as a simple sum. The criterion to be minimized is closely related to the criterion for trait vectors, fixed hypothetical loadings and is designated $\tilde{\Psi}_f$ which is a revision of the criterion in equation (12.52).

$$\tilde{\Psi}_f = \sum_{m=1}^{r} \left\{ \sum_{i=1}^{n} (b_{im} - h_{im})^2 \right\} \quad . \tag{12.129}$$

This minimization is to be constrained by the condition of equation (12.124) which is implemented by defining a function $\tilde{\Psi}_\lambda$:

$$\tilde{\Psi}_\lambda = \sum_{k=1}^{r} \sum_{m=1}^{r} \lambda_{km} (\phi_{km} - \delta_{km}) \tag{12.130}$$

where the λ_{km}'s are Lagrange multipliers and the δ_{km}'s are Kronecker deltas with:

$$\delta_{kk} = 1 \; ; \; \delta_{km} = 0 \quad \text{for } k \neq m \quad . \tag{12.131}$$

The combined criterion is:

$$\tilde{\Psi}_t = \tilde{\Psi}_f + \tilde{\Psi}_\lambda \quad . \tag{12.132}$$

To obtain a minimum of $\tilde{\Psi}_t$ the partial derivative of $\tilde{\Psi}_t$ with respect to element t_{gj} of matrix T is set equal to zero.

$$\frac{\partial\tilde{\Psi}_t}{\partial t_{gj}} = 2\sum_{i=1}^{n}(b_{ig} - h_{ig})a_{ij} + \sum_{m=1}^{r}\lambda_{gm}(t_{mj}) + \sum_{k=1}^{r}\lambda_{kg}(t_{kj}) = 0 \quad . \qquad (12.133)$$

With substitution for b_{jg} from equation (12.128), this equation may be written in matrix form for all entries in T as:

$$TA'A - H'A + \tfrac{1}{2}(\Lambda + \Lambda')T = 0 \quad . \qquad (12.134)$$

Postmultiply by T', note from equation (12.125 that $TT' = I$, and define:

$$P = A'A = P' \quad ; \qquad (12.135)$$

$$S = A'H \quad ; \qquad (12.136)$$

$$Q = \tfrac{1}{2}(\Lambda + \Lambda') = Q' \quad . \qquad (12.137)$$

Then, equation (12.134) may be written as:

$$TPT' + Q = S'T' = TS \quad . \qquad (12.138)$$

The last part of this equation is written since both P and Q are symetric. The solution is facilitated using an Eckart and Young decomposition of matrix S (see Johnsom(1963) for the form used here)

$$S = WDV' \qquad (12.139)$$

where W and V are $r \times r$ orthonormal marices and D is a diagonal matrix. This decomposition is also termed a singular value solution. Then:

$$TS = TWDV'; \qquad (12.140)$$

$$S'T' = VDW'T' \quad .$$

With equation (12.138):

$$TWDV' = VDW'T'$$

and

$$WDV'T = T'VDW' \quad .$$

An equivalence is obtained by setting:

$$T'V = W$$

which results in:

$$T' = WV' \quad . \qquad (12.141)$$

An interesting comparison is between equations (12.139) and (12.141). Diagonal matrix D is eliminated from equation (12.139) and matrix S becomes matrix T'.

To complete the solution for T an eigen solution is obtained for the product matrix $S'S$ which yields:

$$S'S = VD^2V' \quad . \qquad (12.142)$$

With equation (12.139):

$$W = SVD^{-1} \quad . \tag{12.143}$$

A similar solution may be obtained using the product matrix $S\,S'$. There remains a problem as to the algebraic sign to be used in going from D^2 to D and D^{-1}. From equation (12.129):

$$\tilde{\Psi}_f = \sum_{m=1}^{r} \left\{ \sum_{i=1}^{n} b_{im}^2 - 2 \sum_{i=1}^{n} b_{im} h_{im} + \sum_{i=1}^{n} h_{im}^2 \right\} \tag{12.144}$$

In matrix form:

$$\tilde{\Psi}_f = tr(B'B) - 2tr(B'H) + tr(H'H) \quad . \tag{12.145}$$

With the orthogonal transformation:

$$tr(B'B) = tr(A'A) \quad . \tag{12.146}$$

With equation (12.126), (12.136), (12.140), and (12.141):

$$tr(B'H) = tr(TA'H) = tr(TS) = tr(TWDV') = tr(VW'WDV')$$

from which:

$$tr(B'H) = tr(D) \quad . \tag{12.147}$$

Then, with equations (12.146) and (12.147), the value of the criterion from equation (12.145) becomes:

$$\tilde{\Psi}_f = tr(A'A) + tr(H'H) - 2tr(D) \quad . \tag{12.148}$$

With $tr(A'A)$ being the sum of squares of the entries in matrix A and $tr(H'H)$ being the sum of squares in matrix H, both of which are given and positive, criterion $\tilde{\Psi}_f$ is made as small as possible by making $tr(D)$ as large as possible. This is accomplished by choosing the positive square roots of D^2 from equation (12.142).

While the preceding development was for a completely specified hypothetical loadings matrix, Brown (1972) considered the problem of orthogonal rotation to a partially specified target in which not every element in each column of hypothetical loadings needed to be specified. As in the section for Trait Vectors, Fixed Hypothetical Loadings, the set of specified elements in column j of hypothetical loadings will be designated by I_j. The overall criterion considered here is the same as equation (12.52) with the constraint that the trait vectors be orthogonal.

$$\Psi_f = \sum_{m=1}^{r} \left\{ \sum_{i \in I_m} (b_{im} - h_{im})^2 \right\} \quad . \tag{12.52}$$

Instead of a general solution, Browne presented a procedure involving a series of two dimensional rotations for pairs of factors. There was a series of cycles with each cycle involving rotation for every pair of factors. However, before sarting these rotations, Browne presented a system for ordering the factors in matrix A and of reflection of these factors to a most advantageous situation. In the present discussion, the matter of interchanging factors in each pair and the reflection of these factors will be considered before a rotation for the factors in the pair instead of the general reordering of the initial factors. The following material will concentrate on the treatment for a given pair of factors, j and k .

The treatment of factor pair j and k starts from trial loadings b_{ij} and b_{ik} , progresses through tentative loadings $\overset{\#}{b}_{ij}$ and $\overset{\#}{b}_{ik}$ to transformed loadings $\tilde{b}_{ij}$ and $\tilde{b}_{ik}$. The first step from trial loadings to tentative loadings involves possible interchange of the two factors and reflection of the factors. The second step from tentative loadings to transformed loadings invloves the orthogonal rotation of the tentative loadings. For the procedure for interchange and reflection, four coefficients delta are defined:

$$\delta_j = \pm 1 \ ; \quad \delta_k = \pm 1 \ ; \quad \delta_j^* = \pm 1 \quad \delta_k^* = \pm 1 \quad .$$

Two conditions are considered: first that the factors will not be interchanged and second that the factors will be interchanged. Whichever of these conditions apply will be determined later. Following are definitions for each of these conditions.

Without interchange:

$$\overset{\#}{b}_{ij} = \delta_j b_{ij} \quad ; \text{ and } \quad \overset{\#}{b}_{ik} = \delta_k b_{ik} \quad . \tag{12.149}$$

With interchange:

$$\overset{\#}{b}_{ij} = \delta_j^* b_{ik} \quad ; \text{ and } \quad \overset{\#}{b}_{ik} = \delta_k^* b_{ij} \quad . \tag{12.150}$$

Following is the criterion for the tentative loadings for factor pair j and k :

$$\overset{\#}{\Psi}_{fjk} = \sum_{i \in I_j} (\overset{\#}{b}_{ij} - h_{ij})^2 + \sum_{i \in I_k} (\overset{\#}{b}_{ik} - h_{ik})^2 \quad . \tag{12.151}$$

The criterion $\overset{\#}{\Psi}_{fjk}$ is written without reflection and with reflection as follows. Note that the squared deltas equal +1 .

Without reflection:

$$\begin{aligned} \overset{\#}{\Psi}_{fjk} &= \sum_{i \in I_j} (\delta_j b_{ij} - h_{ij})^2 + \sum_{i \in I_k} (\delta_k b_{ik} - h_{ik})^2 \\ &= \sum_{i \in I_j} b_{ij}^2 - 2\delta_j \sum_{i \in I_j} b_{ij} h_{ij} + \sum_{i \in I_j} h_{ij}^2 + \sum_{i \in I_k} b_{ik}^2 - 2\delta_k \sum_{i \in I_k} b_{ik} h_{ik} + \sum_{i \in I_k} h_{ik}^2 \quad . \end{aligned} \tag{12.152}$$

With reflection:

$$\begin{aligned} \overset{\#}{\Psi}{}^*_{fjk} &= \sum_{i \in I_j} (\delta_j^* b_{ik} - h_{ij})^2 + \sum_{i \in I_k} (\delta_k^* b_{ij} - h_{ik})^2 \\ &= \sum_{i \in I_j} b_{ik}^2 - 2\delta_j^* \sum_{i \in I_j} b_{ik} h_{ij} + \sum_{i \in I_j} h_{ij}^2 + \sum_{i \in I_k} b_{ij}^2 - 2\delta_k^* \sum_{i \in I_k} b_{ij} h_{ik} + \sum_{i \in I_k} h_{ik}^2 \quad . \end{aligned} \tag{12.153}$$

In order to make criteria $\overset{\#}{\Psi}_{fjk}$ and $\overset{\#}{\Psi}{}^*_{fjk}$ as small as possible the algebraic signs of the deltas are selected as follow:

$$\text{sign}(\delta_j) = \text{sign}\left(\sum_{i \in I_j} b_{ij} h_{ij}\right) \quad ; \quad \text{sign}(\delta_k) = \text{sign}\left(\sum_{i \in I_k} b_{ik} h_{ik}\right) \quad ; \tag{12.154}$$

$$\text{sign}(\delta_j^*) = \text{sign}\left(\sum_{i \in I_j} b_{ik} h_{ij}\right) \quad ; \quad \text{sign}(\delta_k^*) = \text{sign}\left(\sum_{i \in I_k} b_{ij} h_{ik}\right) \quad . \tag{12.155}$$

Whether to interchange or not to interchange depends on selecting the lesser of criteria $\overset{\#}{\Psi}_{fjk}$ and $\overset{\#}{\Psi}{}^*_{fjk}$:

$$\text{if } \overset{\#}{\Psi}_{fjk} \le \overset{\#}{\Psi}{}^*_{fjk} \text{, do not interchange and use equation (12.149) for the } \overset{\#}{b}\text{'s ;} \tag{12.156}$$

$$\text{if } \overset{\#}{\Psi}_{fjk} > \overset{\#}{\Psi}{}^*_{fjk} \text{, interchange the factors and use equation (12.150) for the } \overset{\#}{b}\text{'s .} \tag{12.157}$$

The transformation from the tentative loadings to the transformaed loadings is performed by a two dimensional orthogonal rotation illustrated below.

$$\begin{bmatrix} \overset{\#}{b}_{1j} & \overset{\#}{b}_{1k} \\ \overset{\#}{b}_{2j} & \overset{\#}{b}_{2k} \\ \vdots & \vdots \\ \overset{\#}{b}_{nj} & \overset{\#}{b}_{nk} \end{bmatrix} \begin{bmatrix} c & -s \\ s & c \end{bmatrix} = \begin{bmatrix} \tilde{b}_{1j} & \tilde{b}_{1k} \\ \tilde{b}_{2j} & \tilde{b}_{2k} \\ \vdots & \vdots \\ \tilde{b}_{nj} & \tilde{b}_{nk} \end{bmatrix}$$

In this transformation c is the cosine of the angle of rotation and s is the sine of the angle of rotation. In equation form:

$$\tilde{b}_{ij} = c\overset{\#}{b}_{ij} + s\overset{\#}{b}_{ik} \quad ; \tag{12.158}$$

$$\tilde{b}_{ik} = -s\overset{\#}{b}_{ij} + c\overset{\#}{b}_{ik} \quad . \tag{12.159}$$

Coefficient c may be restricted to a positive value with the relation between c and s being:

$$c^2 + s^2 = 1 \quad . \tag{12.160}$$

Coefficient c may be considered to be dependent on s, using the positive square root:

$$c = \sqrt{1 - s^2} \quad . \tag{12.161}$$

The criterion for this transformation is:

$$\tilde{\Psi}_{fjk} = \sum_{i \in I_j} (\tilde{b}_{ij} - h_{ij})^2 + \sum_{i \in I_k} (\tilde{b}_{ik} - h_{ik})^2 \tag{12.162}$$

which may be written, with substitution from the preceding equations, as:

$$\tilde{\Psi}_{fjk} = a_0 + a_1 s^2 + a_2 cs + a_3 c + a_4 s \tag{12.163}$$

where:

$$a_0 = \left\{ \sum_{i \in I_j} \overset{\#}{b}{}^2_{ij} + \sum_{i \in I_j} h^2_{ij} + \sum_{i \in I_k} \overset{\#}{b}{}^2_{ik} + \sum_{i \in I_k} h^2_{ik} \right\} \quad ; \tag{12.164}$$

$$a_1 = \left\{ \sum_{i \in I_j} \overset{\#}{b}{}^2_{ik} + \sum_{i \in I_k} \overset{\#}{b}{}^2_{ij} - \sum_{i \in I_j} \overset{\#}{b}{}^2_{ij} - \sum_{i \in I_k} \overset{\#}{b}{}^2_{ik} \right\} \quad ; \tag{12.165}$$

$$a_2 = 2 \left\{ \sum_{i \in I_j} \overset{\#}{b}_{ij} \overset{\#}{b}_{ik} - \sum_{i \in I_k} \overset{\#}{b}_{ij} \overset{\#}{b}_{ik} \right\} \quad ; \tag{12.166}$$

$$a_3 = -2 \left\{ \sum_{i \in I_j} \overset{\#}{b}_{ij} h_{ij} + \sum_{i \in I_k} \overset{\#}{b}_{ik} h_{ik} \right\} \quad ; \tag{12.167}$$

$$a_4 = 2 \left\{ \sum_{i \in I_k} \overset{\#}{b}_{ij} h_{ik} - \sum_{i \in I_j} \overset{\#}{b}_{ik} h_{ij} \right\} \quad . \tag{12.168}$$

Browne (1972) recommends an iterative solution using Bailey's modification of the Newton-Raphson method (see: McCalla, 1967, p. 90). With s being considered as the independent variable, this method involves the first three derivatives of $\tilde{\Psi}_{fjk}$ with respect to s. Note for these derivatives that:

$$\frac{dc}{ds} = \frac{-s}{\sqrt{1 - s^2}} = \frac{-s}{c} \quad . \tag{12.169}$$

Then:

$$\Psi' = \frac{d\tilde{\Psi}_{fjk}}{ds} = 2a_1 s - a_2 \frac{(1 - 2c^2)}{c} - a_3 \frac{s}{c} + a_4 \tag{12.170}$$

$$\Psi'' = \frac{d\Psi'}{ds} = 2a_1 - a_2 s \frac{(1 + 2c^2)}{c^3} - \frac{a_3}{c^3} \quad . \tag{12.171}$$

$$\Psi''' = \frac{d\Psi''}{ds} = -3 \frac{(a_2 + a_3 s)}{c^5} \quad . \tag{12.172}$$

Given trial t of s, s_t, the value of s for trial $t+1$ is given by:

$$s_{t+1} = s_t - \frac{\Psi'}{\left\{ \Psi'' - \Psi'\Psi''' / 2\Psi'' \right\}} \quad . \tag{12.173}$$

The trials are continued until there is a minimal change from s_t to s_{t+1}.

As indicated earlier, the computations involves a series of cycles with each cycle involving a transformation for each pair of factors. Factor j ranges from 1 to $r-1$ with k ranging for each j from $j+1$ to r. The computations are stopped when there are no changes for all pairs of factors in a cycle.

There are two special cases of interest. The first case is when all specified hypothetical loadings are zero, a case closely related to hyperplane fitting. In this case coefficients a_3 and a_4 equal zero. Note that this eliminates the last two terms from equation (12.162) for the criterion and from the first derivative in equation (12.170). Note, also, the changes in the equations for the second and third derivatives in equations (12.171) and 912.172). From the interchange and reflection portion of transformation for a pair of factors, the reflections of factors are indeterminant since all terms $(\sum bh)$ equal zero. A further feature after the interchange: $a_1 \geq 0$.

There are two special situations as follow.

When $a_2 = 0$, coefficients s and c may be set as:

$s = 0$ and $c = 1$.

When $a_1 = 0$ and $|a_2| > 0$ (assume, as will be developed, that $|c| > 0$) from the first derivative, equation (12.170):

$1 - 2c^2 = 0$

so that:

$c = \sqrt{1/2}$

$s = \pm\sqrt{1/2}$

with the sign of s set so that the term $a_2 s \frac{(1 + 2c^2)}{c^3}$ in the second derivative is negative so as to make the contribution of this term positive to the second derivative; that is:

$\text{sign}(s) = \text{sign}(-a_2)$.

The general solution is considered using two conditions to control the sizes of numbers in the computations.

When $|a_1| \geq |a_2| > 0$, define, in consideration of setting the first derivative, equation (12.170), equal to zero:

$$x = -\frac{|a_2|}{|a_1|} = \frac{2\delta^* cs}{(1 - 2c^2)} \tag{12.174}$$

with $\delta^* = \pm 1$, the sign being selected such that the term $a_2 s \frac{(1 + 2c^2)}{c^3}$ in the second derivative is negative so as to make a positive contribution to the second derivative. Then:

$1 + x^2 = 1\Big/(1 - 2c^2)^2$

or, using the positive square root and $\delta = \pm$:1

$(1 - 2c^2) = \delta\Big/\sqrt{1 + x^2}$.

Then:

$c^2 = \frac{1}{2}\left(1 + \delta\Big/\sqrt{1 + x^2}\right)$.

For $c^2 \geq \frac{1}{2}$, $\delta = +1$ so that, using a positive square root:

$c = \sqrt{\frac{1}{2}\left(1 + 1\Big/\sqrt{1 + x^2}\right)}$.

From equation (12.174):

$s = \delta^* x\Big/2c\sqrt{1 + x^2}$

with the algebraic sign of δ^* being:

$\text{sign}(\delta^*) = \text{sign}(a_2)$.

When $|a_2| > |a_1| > 0$, define:

$$y = -\frac{|a_1|}{|a_2|} = \frac{\delta^*(1 - 2c^2)}{2cs} \tag{12.175}$$

with $\delta^* \pm 1$, the sign being selected such that the term $a_2 s \frac{(1 + 2c^2)}{c^3}$ in the second derivative is negative so as to make the contribution positive to the second derivative. Then:

$1 + y^2 = 1\Big/4c^2 s^2$

and

$\frac{y^2}{1 + y^2} = (1 - 2c^2)^2$.

Using the positive square root and a coefficient $\delta = \pm 1$:

$$1 - 2c^2 = \delta\sqrt{\frac{y^2}{1+y^2}} \qquad (12.176)$$

so that:

$$c^2 = \frac{1}{2}\left[1 - \delta\sqrt{\frac{y^2}{1+y^2}}\right] \quad .$$

In order for $c^2 \geq \frac{1}{2}$, coefficient δ is set to equal -1 . The, using the positive square root:

$$c = \sqrt{\frac{1}{2}\left[1 + \sqrt{\frac{y^2}{1+y^2}}\right]}$$

From equations (12.175) and (12.176):

$$s = \delta^* \Big/ 2c\sqrt{1+y^2} \quad .$$

As indicated above, the algebraic sign of δ^* is to be set as follows:

$$\mathrm{sign}(\delta^*) = \mathrm{sign}(-a_2) \quad .$$

The second special case occurs when all the specified hypothetical loadings for factors j and k are for the same entities; that is when $I_j = I_k$. The most collmon situation is that the matrix of hypothetical loadings is completely specified. In this case coefficients a_1 and a_2 equal zero. then, from equations (12.163), (12.170), (12.171):

$$\tilde{\Psi} = a_0 + a_3 c + a_4 s \quad ;$$

$$\Psi' = -a_3 \frac{s}{c} + a_4 \quad ;$$

$$\Psi'' = -\frac{a_3}{c^3} \quad .$$

After setting the first derivative equal to zero algebraic operations yield:

$$\frac{s}{c} = \frac{a_4}{a_3}$$

$$c^2 = \frac{a_3^2}{a_3^2 + a_4^2} \quad .$$

Then, using the positive square root and $\delta = \pm 1$:

$$c = \frac{\delta\, a_3}{\sqrt{a_3^2 + a_4^2}} \quad ;$$

$$s = \frac{\delta\, a_4}{\sqrt{a_3^2 + a_4^2}} \quad .$$

With the interchange and reflections a_3 is negative, see equation (12.167), so that for the second derivative to be positive, as is necessary for a minimum solution:

$$\delta = -1 \quad .$$

All of the preceding solutions yield a matrix B but not the transformation matrix T . In order to obtain the matrix T , a solution from equation (12.126) yields:

$$T' = (A'A)^{-1}(A'B) \quad .$$

Table 12.16 gives the results for orthogonal trait vectors, fixed hypothetical loadings for the fourth year medical students data. As in previous tables, the matrix of hypothetical loadings is given on the left with the transformation results being given on the right. These results are to

Table 12.16
Solution For: Orthogonal Trait Vectors, Fixed Hypothetical Loadings
Fourth Year Medical Students Data

<u>Hypothetical Loadings</u>

<u>Transformatin Results</u>

Trait Vectors

	1	2
1	.635	.772
2	.772	-.635

Given Loadings

	1	2
1	.568	.090
2	.486	-.113
3	.393	.005
4	-.005	.844
5	.013	.852

Factor Loadings

	1	2
1	.504	.278
2	.688	.105
3	.697	.141
4	.183	.847
5	.171	.829

Sums of Squared Differences

1	2
.198	.102

Criterion ψ_f = .300

be compared with those given in Table 12.12 when the trait vectors were permitted to be oblique. Most notable for the factor loadings is that the small loadings, now, are measurably positive rather than being near zero, some of the previous loadings being negative. This is the result of the shift in the transformed factors from having a correlation of .360 to the restricted correlation of zero. Further, the sums of squared differences have been increased with the criterion Ψ increasing from .175 for the oblique case to .300 for the orthogonal case. A question to be considered is whether the loss of goodness of fit is worth the reduction in number of parameters due to the restriction on the correlation between the factors.

12.2.2.2 Orthogonal Trait Vectors, Proportional Hypothetical Loadings

The option of using proportional hypothetical loadings is available when the analyst consideres that the hypothetical loadings are given only within constants of proportionality, one such constant for each factor. However, the trait vectors are to be orthogonal. A restriction is that at least one of the specified hypothetical loadings for each factor must be greater than zero in absolute value. The criterion for this case is the same as for the case permitting oblique factors:

$$\tilde{\Psi}_p = \sum_{m=1}^{r} \left\{ \sum_{i \in I_m} (b_{im} - c_m h_{im})^2 \right\} \quad . \tag{12.76}$$

where c_m is a constant of proportionality for each factor applied to the hypothetical loadings for that factor. However, now, this criterion is to be minimized under the constraint of equation (12.124) that the trait vectors be orthogonal. This constrain is implemented by the definition of a function $\tilde{\Psi}_\lambda$ defined in equation (12.130)

$$\tilde{\Psi}_\lambda = \sum_{k=1}^{r} \sum_{m=1}^{r} \lambda_{km}(\phi_{km} - \delta_{km}) \tag{12.130}$$

where the λ_{km}'s are Lagrange multipliers and the δ_{km}'s are Kronecker deltas defined in equations (12.131):

$$\delta_{kk} = 1 \ ; \quad \delta_{km} = 0 \quad \text{for } k \neq m \quad . \tag{12.131}$$

The combined criterion is:

$$\tilde{\Psi}_t = \tilde{\Psi}_p + \tilde{\Psi}_\lambda \quad . \tag{12.132}$$

Solutions for the c_m's are considered first for given trail trait vectors and loadings in matrix B . The partial derivative of $\tilde{\Psi}_t$ with respect to the c_m is set equal to zero.

$$\frac{\partial \tilde{\Psi}_t}{\partial c_m} = 2 \sum_{i \in I_m} (b_{im} - c_m h_{im})\, h_{im} = 0 \quad . \tag{12.177}$$

This yields:

$$c_m = \left(\sum_{i \in I_m} b_{im} h_{im} \right) \Big/ \left(\sum_{i \in I_m} h_{im}^2 \right) \quad . \tag{12.178}$$

Adjusted hypothetical loadings are obtained by:

$$\ddot{h}_{im} = c_m h_{im} \quad . \qquad (12.179)$$

An alternating type of solution is utilized starting from initial trial coefficients c_m , possibly using all c_m equal to unity. Adjusted hypothetical loadings are computed by equation (12.179) and initial trial loadings in matrix B are determined using the adjusted hypothetical loadings as if they were fixed and the solution for orthogonal trait vectors, fixed hypothetical loadings in the preceding section. Having the trial loadings in matrix B , new trial coefficients c_m are computed by equation (12.179). New adjusted hypothetical loadings are computed by equation (12.179) and new trial loadings in matrix B are computed using the adjusted hypothetical loadings as if they are fixed and the solution for orthogonal trait vectors, fixed hypothetical loadings. This alternation between trial c_m's and trial matrices B is continued until there is a minimal change in the constants of proportionality, c_m, from trial to trial.

Table 12.17 presents the results for this criterion for the fourth year mdeical students. As before, the hypothetical loadings are given onthe left with the given hypothetical loadings at the far left. In the second section for the hypothetical loadings are the obtained constants of proportionality and the adjusted hypothetical loadings. The transformation results are given at the right with the trait vectors, factor loadings, and sums of squared differences between the factor loadings and the adjusted hypothetical loadings. A comparison of these results given in Table 12.13 for the case when the trait vectors were permitted to be oblique indicates that the constants of proportionality in these two tables are very similar. The first factor hypothetical loadings are adjusted by a constant more than 1.25 while the constants of proportionality for the second factor are slightly lexx than unity. Again, as for the preceding section in Table 12.16, the factor loadings for small given hypothetical loadings are markedly positive. The sums of squared differences for the orthogonal trait vectors, proportional hypothetical loadings are less than for the case in Table 12.16 for fixed hypothetical loadings. However, these sums of squared differences are greater than for the case in Table 12.13 when the trait vectors were permitted to be oblique. Again, there is a question for the fourth year medical students data whether or not the restriction for orthogonal trait vectors is more desirable than the case when the trait vectors were permitted to be oblique.

12.3 Factor Analysis of Several Batteries in One Sample

Two major purposes are served by factor analyses of several batteries in one sample. The firstn purpose is to confirm generalization of factors from one battery to a second battery. Tucker (1958) introduced inter-battery factor analysis for the purpose of generalizing factors across

Table 12.17

Solution For: Orthogonal Trait Vector, Proportional Hypothetical Loadings
Fourth Year Medical Students Data

Hypothetical Loadings | Transformation Results

Constants of Proportionality

1	2
1.267	.977

Trait Vectors

	1	2
1	.659	.752
2	.752	-.659

Given Loadings

	1	2
1	.568	.090
2	.486	-.113
3	.393	.005
4	-.005	.844
5	.013	.852

Adjusted Loadings

	1	2
1	.720	.088
2	.615	-.110
3	.498	.005
4	-.006	.825
5	.016	.833

Factor Loadings

	1	2
1	.512	.263
2	.691	.083
3	.701	.119
4	.209	.841
5	.196	.824

Sums of Squared Differences

1	2
.169	.081

Criterion ψ_p = .250

selections of attributes measures in contrast to generalization of results for a given battery over samplinsg of individuals. The second purpose is to determine factors which are common to several batteries separate from factors which appear in single batterise. For an example consider one battery of course grades in secondary school, a second battery of entrance examinations to college, and a third battery of college grades. Here the questions could involve what factors are common to pairs of these batteries and are common through all three of the batteries. For an example, there may be factors common to secondary school grades and college grades which are not involved in the entrance examinations.

The model for inter-battery factor analysis is to be considered first followed by discussion of the generalized model for three or more batteries. Equation (12.180) gives the model in the population for two batteries where Σ is the covariance matrix for all attributes. There are n_1 attributes in battery 1 and n_2 attributes in battery 2 . This matrix is partitioned into sections for the intercovariances for each battery and covariance matrices between the two batteries.

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} = \begin{bmatrix} \mathbb{A}_1 & \mathbb{S}_1 & 0 \\ \mathbb{A}_2 & 0 & \mathbb{S}_2 \end{bmatrix} \begin{bmatrix} \mathbb{A}_1' & \mathbb{A}_2' \\ \mathbb{S}_1' & 0 \\ 0 & \mathbb{S}_2' \end{bmatrix} \qquad (12.180)$$

Matrices $\mathbb{A}_1$ and $\mathbb{A}_2$ are the common factor matrices for factors common to the two batteries while matrices $\mathbb{S}_1$ and $\mathbb{S}_2$ contain the factors specific to each of the two matrices. Note that the attribute specific factors are included in matrices $\mathbb{S}_1$ and $\mathbb{S}_2$. There are r inter-battery common factors. In a sample, the model for covariance matrix C is:

$$C = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} = \begin{bmatrix} A_1 & S_1 & 0 \\ A_2 & 0 & S_2 \end{bmatrix} \begin{bmatrix} A_1' & A_2' \\ S_1' & 0 \\ o & S_2' \end{bmatrix} + \begin{bmatrix} \Delta_{11} & \Delta_{12} \Delta_{12} \\ \Delta_{21} & \Delta_{22} \end{bmatrix} \qquad (12.181)$$

where matrices A_1 and A_2 contain the inter-battery common factors in the sample while matrics S_1 and S_2 contain the battery specific factors in the sample. Matrix Δ contains the discrepancies of fit. An interesting feature of this model is that matrices C_{12} and C_{21} do not involve the battery specific factors.

$$C_{12} = A_1 A_2' + \Delta_{12} = C_{21}' \quad . \qquad (12.182)$$

Tucker (1958) proposed a least squares solution for the inter-battery factors using an Eckart and Young (1936) decomposition (This is the same as a singular value solution.). Equation (12.183) symbolizes this solution.

$$C_{12} = V_1 D V_2' = \begin{bmatrix} V_{1r} & V_{1\delta} & V_{10} \end{bmatrix} \begin{bmatrix} D_r & 0 & 0 \\ 0 & D_\delta & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} V_{2r}' \\ V_{2\delta}' \\ V_{20}' \end{bmatrix} \qquad (12.183)$$

where V_1 and V_2 are orthonormal matrices, V_1 being n_1 x n_1 and V_2 being n_2 x n_2 . Matrix D has diagonal sections D_r and D_δ . There are r columns in matrix sections V_{1r} and V_{2r} with r being the number inter-battery factors retained in the analysis. D_r has r entries.

Let there be R non-zero entries in the diagonal of D, then the number of columns in $V_{1\delta}$, $V_{2\delta}$ are $R - r$ and the number of entries in D_δ is $R - r$. The value chosen for r is discussed below. The direction of the column vectors in V_1 and V_2 are to be chosen such that the diagonal entries in D_r and D_δ are positive. Initial matrices A_1 and A_2 are given by:

$$A_1 = V_{1r}D_r^{1/2} \quad ; \tag{12.184.1}$$

$$A_2 = V_{2r}D_r^{1/2} \quad . \tag{12.184.1}$$

The diecrepancy matrix Δ_{12} is given by:

$$\Delta_{12} = V_{1\delta}D_\delta V'_{2\delta} \quad . \tag{12.185}$$

The sum of squares of entries in Δ_{12} is given by:

$$SSQ(\Delta_{12}) = \sum_{j=r+1}^{R} d_j^2 \quad . \tag{12.186}$$

An alternative is:

$$SSQ(\Delta_{12}) = SSQ(C_{12}) - \sum_{j=1}^{r} d_j^2 \quad . \tag{12.187}$$

The series of entries d_j^2 , j = 1,R , may be inspected to determine the number of inter-battery factors, r, such that the last dimension accepted has the last d_j^2 of acceptable size.

The initial matrices A_1 and A_2 determined in equations (12.184.1) and (12.184.2) are subject to a general transformation. Let T , r x r , be a general, nonsingular transformation matrix. Then transformed inter-battery matrices, $\tilde{A}_1$ and $\tilde{A}_2$, are obtained by:

$$\tilde{A}_1 = A_1 T \quad ; \tag{12.188.1}$$

$$\tilde{A}_2 = A_2(T')^{-1} \quad . \tag{12.188.2}$$

Then, $\tilde{A}_1$ and $\tilde{A}_2$ may be substituted in equation (12.182) for A_1 and A_2 . Note that this transformation is more geneneral than those usually encountered in factor analysis, there is no restriction on the lengths of the column vectors of T . This presents a very difficult problem in the treatment of inter-battery factor analysis.

Browne (1979) described a maximum likelyhood solution in inter-battery factor analysis. He noted, also, the general transformation described in the preceding paragraph.

The inter-battery model may be extended to three or more batteries (the number of batteries being designated by p) as indicated in the following equation for the population.

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} & \cdots & \Sigma_{1p} \\ \Sigma_{12} & \Sigma_{22} & \cdots & \Sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_{p1} & \Sigma_{p2} & \vdots & \Sigma_{pp} \end{bmatrix} = \begin{bmatrix} \mathbb{A}_1 & \mathbb{S}_1 & 0 & \cdots & 0 \\ \mathbb{A}_2 & 0 & \mathbb{S}_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbb{A}_p & 0 & 0 & \cdots & \mathbb{S}_p \end{bmatrix} \cdots \begin{bmatrix} \mathbb{A}'_1 & \mathbb{A}'_2 & \cdots & \mathbb{A}'_p \\ \mathbb{S}'_1 & 0 & \cdots & 0 \\ 0 & \mathbb{S}'_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbb{S}'_p \end{bmatrix}$$

(12.189).