Between-Groups Comparison of Principal Components

W. J. KRZANOWSKI*

A method is given for comparing principal component analyses conducted on the same variables in two different groups of individuals, and an extension to the case of more than two groups is outlined. The technique leads to a latent root and vector problem, which has also arisen in the comparison of factor patterns in separate factor analyses. Emphasis in the present article is on the underlying geometry and interpretation of the results. An illustrative example is provided.

KEY WORDS: Analytical rotation; Congruence coefficient; Latent roots and vectors; Principal component analysis.

1. INTRODUCTION

Principal component analysis (PCA) is a widely used technique in multivariate analysis (see Anderson 1958, Kendall and Stuart 1968, Seal 1964). It may be used on any data matrix in which p variables (responses) are measured on each of n individuals (units). One of its most useful aspects is that of multivariate description, since it identifies those linear combinations of responses (i.e., components) that have most variability in the sample. Such components frequently carry biological or other interpretations that provide valuable insight into the mechanisms generating the data.

Suppose that several groups of individuals have the same variables measured on them, and interest centers on discovering how similar the groups are with respect to their overall features. An intuitively appealing procedure is to describe each group in terms of a small number (we hope) of principal components and assess the similarity of the groups by comparing these components. Unfortunately, visual inspection is not very trustworthy, as two sets of components that are quite different in appearance may in fact define the same subspace of the original multivariate space (defined in Section 2). Thus, some reliable analytical technique is required for the comparison.

A similar problem has bedeviled factor analysts over the years, because of the rotational indeterminacy of a factor analysis solution (Lawley and Maxwell 1971, p. 7). In order to test a hypothesized factor pattern, first the matrix of factor loadings must be transformed to maximum similarity with the target matrix (Horst 1956; Cliff 1962; Hurley and Cattell 1962). In order to compare two factor patterns, first they must be transformed to maximal agreement with each other (Cliff 1966; Schonemann and Carroll 1970; Evans 1971). The main problem throughout is the choice of criterion for judging similarity. Various criteria have been proposed (Korth 1973), but the congruence coefficient has found some favor by virtue of its suitable properties (Korth and Tucker 1976). This was the choice for Cliff's (1966) solution.

The geometrical interpretation of the congruence coefficient between two p-component vectors is simply as the cosine of the angle between them when they are regarded as two lines in p-dimensional Euclidean space. Therefore, the essentially geometrical nature of PCA (Gower 1967) suggests that an analogous technique can be developed for the comparison of principal components. This is done in Section 2. Section 3 then proposes an extension of the technique to the simultaneous comparison of $g(>2)$ groups, and examples are given in Section 4. Similarities of the proposed techniques with other established procedures are noted in Section 5.

2. ANALYSIS FOR TWO GROUPS

Consider two multivariate samples A and B of n_1 and n_2 individuals, respectively, each individual having the same p variables measured on it. Suppose each sample has undergone PCA. Let the original variables be denoted by $x_1, x_2, \ldots, x_p$, and denote the principal components of A and B by $y_1, \ldots, y_p$ and $z_1, z_2, \ldots, z_p$, respectively. Then $y_i = \sum_{j=1}^p l_{ij}x_j$ and $z_i = \sum_{j=1}^p m_{ij}x_j$ ($i = 1, \ldots, p$), where l_{ij} and m_{ij} ($i = 1, \ldots, p$; $j = 1, \ldots, p$) are the principal component loadings in the two groups. It will be assumed that the usual normalizations $\sum_{j=1}^p l_{ij}^2 = \sum_{j=1}^p m_{ij}^2 = 1$ ($i = 1, \ldots, p$) have been adopted. If each of the original variables $x_1, \ldots, x_p$ is identified with an orthogonal axis in p-dimensional Euclidean space, the two groups A and B are represented by two swarms of points in this space. Principal component analysis is simply a rotation of axes to new positions $y_1, \ldots, y_p$ for group A and $z_1, \ldots, z_p$ for group B. These new axes are such that the orthogonal projections of the sample points on them have decreasing spread. In this interpretation, l_{ij} is the direction cosine of the ith component of A with the axis corresponding to x_j, and m_{ij} has a similar interpretation for group B. This geometrical representation has been well documented (e.g., Gower 1967).

Suppose that k components are considered adequate for the purposes of representing each sample, and let $\mathbf{L}$ and $\mathbf{M}$ be matrices with elements l_{ij} and m_{ij}, respectively

* W. J. Krzanowski is a Lecturer, Department of Applied Statistics, University of Reading, Whiteknights, Reading, RG6 2AN, England. The author is indebted to B. Carroll of the British Council for permission to use the data and to some referees for their constructive comments and criticisms of an earlier version of this article.

© Journal of the American Statistical Association
September 1979, Volume 74, Number 367
Theory and Methods Section

$(i = 1, \ldots, k;\ j = 1, \ldots, p)$. Thus the swarm of points from each sample can be considered as effectively embedded in a k-dimensional subspace of the original p-dimensional space, and these two subspaces are defined by the orthogonal axes $y_1, \ldots, y_k$ and $z_1, \ldots, z_k$, respectively. In order to compare the two sets of principal components, it is necessary to compare the two k-dimensional subspaces that they generate. The following theorems provide a means of doing this.

Theorem 1: The minimum angle between an arbitrary vector in the space of the first k principal components of A and the one most nearly parallel to it in the space of the first k components of B is given by $\cos^{-1}\{(\lambda_1)^{\frac{1}{2}}\}$, where λ_1 is the largest eigenvalue of $\mathbf{S} = \mathbf{LM'ML'}$.

Proof: Let $\mathbf{a}$ be an arbitrary (unit) vector with elements $a_1, \ldots, a_k$ in the subspace generated by the k principal components of A. Its coordinates referred to the original p orthogonal axes are therefore the components of $\mathbf{b} = \mathbf{L'a}$. The projection of this vector onto the subspace generated by the first k components of B is $\mathbf{M'Mb}$, so the angle $\delta(0 \le \delta \le \pi)$ between the two satisfies $\cos^2\delta = \mathbf{b'M'Mb} = \mathbf{a'LM'ML'a}$.

Now, minimizing δ is equivalent to maximizing $\cos^2 \delta$. It is well known that $\sup_{||\mathbf{x}||=1} \mathbf{x'Cx}$ is given by the largest eigenvalue λ_1 of $\mathbf{C}$ and is attained when $\mathbf{x}$ is the eigenvector corresponding to λ_1. Thus the value of $\mathbf{a}$ that minimizes δ is given by the eigenvector corresponding to the largest eigenvalue λ_1 of $\mathbf{S} = \mathbf{LM'ML'}$, and $\cos^2\delta = \lambda_1$ for this value of $\mathbf{a}$. It can be shown easily that all eigenvalues λ of $\mathbf{S}$ satisfy $0 \le \lambda \le 1$, so the result follows directly.

Theorem 2: Let λ_i be the ith largest eigenvalue of $\mathbf{S}$, $\mathbf{a}_i$ its associated eigenvector, and $\mathbf{b}_i = \mathbf{L'a}_i$ $(i = 1, \ldots, k)$. Then $\mathbf{b}_1, \ldots, \mathbf{b}_k$ form a set of mutually orthogonal vectors embedded in subspace A and $\mathbf{M'Mb}_1, \ldots, \mathbf{M'Mb}_k$, a corresponding set of mutually orthogonal vectors in subspace B into which the differences between the subspaces can be partitioned. The angle between the ith pair $\mathbf{b}_i$, $\mathbf{M'Mb}_i$ is given by $\cos^{-1}(\lambda_i)^{\frac{1}{2}}$ $(i = 1, \ldots, k)$.

Proof: $\mathbf{b}_i'\mathbf{b}_j = \mathbf{a}_i'\mathbf{LL'a}_j = \mathbf{a}_i'\mathbf{a}_j$ and $(\mathbf{M'Mb}_i)'(\mathbf{M'Mb}_j) = \mathbf{b}_i'\mathbf{MM'b}_j = \mathbf{a}_i'\mathbf{Sa}_j = \lambda_j\mathbf{a}_i'\mathbf{a}_j$ (since $\mathbf{LL'} = \mathbf{MM'} = \mathbf{I}$ and $\mathbf{Sa}_j = \lambda_j\mathbf{a}_j$). Orthogonality of the two sets of vectors thus follows from orthogonality of the $\mathbf{a}_j$.

It has been shown in Theorem 1 that $\mathbf{b}_1$ and $\mathbf{M'Mb}_1$ give the two closest vectors in the original space when one is constrained to be in subspace A and the other in subspace B. Continuing the eigenvalue and eigenvector decomposition of $\mathbf{S}$, it follows that $\mathbf{b}_2$ and $\mathbf{M'Mb}_2$ give directions, orthogonal to the previous ones, along which the next smallest angle between the subspaces is represented. An argument analogous to that of Theorem 1 gives the angle between these two vectors as $\cos^{-1}(\lambda_2)^{\frac{1}{2}}$. Completing the decomposition yields the required result.

Let θ_{ij} be the angle between the ith principal component of group A and the jth principal component of group B. Then $\cos\theta_{ij}$ is the (i, j)th element of $\mathbf{T} = \mathbf{LM'}$ so that

$$\sum_{i=1}^{k} \lambda_i = \operatorname{trace} \mathbf{S} = \operatorname{trace} \mathbf{TT'} = \sum_{i=1}^{k}\sum_{j=1}^{k} \cos^2\theta_{ij} .$$

Thus the sum of the eigenvalues of $\mathbf{S}$ equals the sum of squares of the cosines of the angles between each of the k eigenvectors defining the principal components of A and each one of B. This sum can be used as a measure of total similarity between the two spaces. The value of the sum is easily seen to lie between k (coincident spaces) and 0 (orthogonal spaces). Theorems 1 and 2 then show that the similarities between A and B can be exhibited solely through the pairs $\mathbf{b}_i$, $\mathbf{M'Mb}_i$, λ_i measuring the contribution of the ith pair to the total similarity.

Corollary 1: If the two subspaces defined by the principal components of A and B intersect in an r-dimensional subspace of the original p-dimensional space, then the first r eigenvalues of $\mathbf{S}$ have value 1, and $\mathbf{b}_1, \ldots, \mathbf{b}_r$ form a basis for this subspace.

Corollary 2: If A and B have been characterized by k_1 and k_2 components, respectively $(k_1 \ne k_2)$, then $\mathbf{T}$ is a $(k_1 \times k_2)$ matrix of rank $k = \min(k_1, k_2)$. Hence $|k_1 - k_2|$ zero eigenvalues will always exist for the larger of $\mathbf{T'T}$ and $\mathbf{TT'}$.

Comparison between A and B can proceed on the basis of the k nonzero eigenvalues and corresponding eigenvectors in such a case.

These results thus furnish some descriptive tools for the comparison of A and B. Of course, descriptive tools are only as good as our ability to interpret and relate them to quantities of interest. In the present analysis, the minimum angles $\cos^{-1}(\lambda_i)^{\frac{1}{2}}$ between subspaces generated by sets of principal components will provide a measure of the extent to which the subspaces differ. In order to interpret the nature of the similarities or differences, we can then consider the pair of eigenvectors $\mathbf{b}_i$ and $\mathbf{M'Mb}_i$ associated with each eigenvalue λ_i. These vectors are defined with respect to the original p axes and hence may be interpreted by reference to the p coefficients in each vector. Thus inspection of the coefficients may reveal, for example, some biological explanation for the vectors that are most alike between the two groups. Furthermore, a natural way of obtaining the single vector in the original p-dimensional space that is closest to both these vectors would be as the bisector of the angle between them, in the plane in which they lie. This bisector is given by

$$\mathbf{c}_i = \{1/(1 + 3\lambda_i)^{\frac{1}{2}}\}(\mathbf{I} + \mathbf{M'M})\mathbf{b}_i .$$

The set $\mathbf{c}_1, \ldots, \mathbf{c}_k$ then defines the overall k-dimensional subspace that is closest to both subspaces A and B.

Finally, some comparison can be made between this technique and Cliff's (1966) for comparing factor analysis solutions. The similarities result from the equivalence of the congruence coefficient and the cosine of the angle between a pair of vectors in multidimensional space. Cliff, however, was mainly concerned with an algebraic transformation that produced sets of factor loadings resembling one another as closely as possible, and his sole aim was to produce these transformed factor patterns. This development for PCA has led to a rather more general comparison. Here the transformation itself plays a more central role, the eigenvalues can be used to assess the degree of resemblance between principal component subspaces, and the eigenvectors can be used to explore the nature of this resemblance.

3. ANALYSIS FOR MORE THAN TWO GROUPS

Suppose that $g(>2)$ groups exist, with n_i units in the ith group ($i = 1, \ldots, g$), and the same p variables measured on each unit. Also, each group has been described by its first k principal components. Write $\mathbf{L}_t$ as the matrix whose (i, j)th element is $l_{ij}^{(t)}$, the loading of the jth variable on the ith principal component for the tth group ($i = 1, \ldots, k$; $j = 1, \ldots, p$; $t = 1, \ldots, g$). We now consider the problem of simultaneously comparing these g sets of components.

Theorem 3: Let $\mathbf{b}$ be an arbitrary vector in the original p-dimensional data space, and let δ_t be the angle between $\mathbf{b}$ and the vector most nearly parallel to it in the space generated by the k principal components of group t ($t = 1, \ldots, g$). Then the value of $\mathbf{b}$ that minimizes $V = \sum_{t=1}^{g} \cos^2 \delta_t$ is given by the eigenvector $\mathbf{b}_1$ corresponding to the largest eigenvalue μ_1 of $\mathbf{H} = \sum_{t=1}^{g} \mathbf{L}_t'\mathbf{L}_t$.

Proof: From the definition of δ_t we see that $\cos^2 \delta_t = \mathbf{b}'\mathbf{L}_t'\mathbf{L}_t\mathbf{b}$, so that $V = \sum_{t=1}^{g} \cos^2 \delta_t = \mathbf{b}'\mathbf{H}\mathbf{b}$, and the theorem follows from the standard eigenvalue-eigenvector result given in the proof of Theorem 1.

Therefore, if we consider V to be a measure of closeness of $\mathbf{b}$ to all the k-dimensional subspaces, then the average component that agrees most closely with all g sets of k principal components is given by $\mathbf{b}_1$. A measure of discrepancy between this component and the k-dimensional subspace for the tth group ($t = 1, \ldots, g$) is given by

$$\delta_t = \cos^{-1} \{ (b_1'\mathbf{L}_t'\mathbf{L}_t\mathbf{b}_1)^{\frac{1}{2}} \} \ . \tag{3.1}$$

Completing the latent root and vector analysis of $\mathbf{H}$ will lead to the subspace of dimension k that resembles all g subspaces as closely as possible. The eigenvector $\mathbf{b}_2$ corresponding to the second-largest eigenvalue μ_2 will give the direction, orthogonal to $\mathbf{b}_1$, in which the criterion V has next largest value. The angle between $\mathbf{b}_2$ and the vector most nearly parallel to it in the subspace for the tth group is found as in (3.1). This procedure can be continued for all k latent roots and vectors, and these k vectors define the required subspace. If k_i components have been obtained for the ith group ($i = 1, \ldots, g$) and $k = \min(k_1, k_2, \ldots, k_g)$, then only a k-dimensional comparison will be useful. Any further dimension will be orthogonal to at least one of the group subspaces.

That this procedure reduces to the solution obtained in Section 2 when $g = 2$ can be demonstrated easily as follows.

The solution of Section 2 required maximization of $\mathbf{a}'\mathbf{L}\mathbf{M}'\mathbf{M}\mathbf{L}'\mathbf{a}$ subject to $\mathbf{a}'\mathbf{a} = 1$. This is equivalent to maximizing $\mathbf{a}'\mathbf{L}(\mathbf{L}'\mathbf{L}+\mathbf{M}'\mathbf{M})\mathbf{L}'\mathbf{a}$ subject to $(\mathbf{a}'\mathbf{L})(\mathbf{L}'\mathbf{a})=1$, since $\mathbf{L}\mathbf{L}' = \mathbf{I}$. Setting $\mathbf{b} = \mathbf{L}'\mathbf{a}$ as in Section 2, it follows that the quadratic form is maximized when $\mathbf{b}$ is the eigenvector associated with the largest eigenvalue of $(\mathbf{L}'\mathbf{L} + \mathbf{M}'\mathbf{M})$. Hence $\mu\mathbf{b} = (\mathbf{L}'\mathbf{L} + \mathbf{M}'\mathbf{M})\mathbf{b} = (\mathbf{I} + \mathbf{M}'\mathbf{M})\mathbf{b}$ on using $\mathbf{b} = \mathbf{L}'\mathbf{a}$. This establishes the equivalence of the vector $\mathbf{b}$ in the two sections.

4. EXAMPLES

The following analyses are abstracted from a larger investigation into the educational achievements of Venezuelan students at colleges in the north of England. The analyses are presented here as an illustration of the foregoing techniques and are not intended to be representative of any conclusions reached in the full investigation.

The data consisted of examination scores achieved in eight subjects by 160 students. These students were distributed among 10 different colleges, which meant that sample sizes were quite small for each, lying between 10 and 20. One aim was to compare performance between colleges, so principal component analyses were done separately on the raw scores from each college in an attempt to summarize descriptively the main sources of variation between students in each. Three colleges, labeled A, B, and C, have been selected for present illustration. The loadings on each of the first four principal components for each college are displayed in Table 1. The first component in each case made up about 60 percent of the trace and the second about 20 percent.

In order to demonstrate the technique of Section 2, the principal components of colleges A and B were compared, and the results are displayed in Table 2. The first r principal components in the two groups are compared sequentially for $r = 1$, 2, 3, and 4. Part (a) of Table 2 gives the angles of separation of the two principal component spaces. The values of the latent roots, λ, are quoted, as well as the corresponding angles, given by $\cos^{-1} (\lambda)^{\frac{1}{2}}$. Part (b) of Table 2 then gives the vectors (referred to the original axes) that are closest to both spaces for each dimensionality of comparison.

A comparison of the principal components for colleges A and B can now be made more readily than by inspecting Table 1. We see that the first principal components in each of the colleges do not compare very well, with an angle of 34° between them (although an unreported Monte Carlo study found that 34° is near the 95th percentile for one common component with $n = 20$). Similarly, when the first two components of each college are compared, the nearest the two spaces come to coincidence is 30°. Otherwise, they are almost orthogonal to each other. When we proceed to component spaces of dimensions 3 and 4, however, we see that a common space of three dimensions is strongly indicated, with

1. Loadings on First Four Principal Components for Each of Three Colleges

	College A				*College B*				*College C*			
Subject	*I*	*II*	*III*	*IV*	*I*	*II*	*III*	*IV*	*I*	*II*	*III*	*IV*
Comprehension	.227	.458	.218	.804	.207	.655	.669	.016	.402	.465	.181	.291
Essay	.322	−.191	.331	.063	.376	.285	−.277	−.009	.405	−.040	.128	−.553
Cloze 1	.319	.004	−.525	.000	.250	−.098	.169	.110	.402	−.193	−.581	.225
Cloze 2	.275	−.014	−.416	−.040	.184	.091	.137	.023	.321	−.209	−.505	.097
Structure	.741	−.123	.385	−.312	.400	.070	−.211	−.839	.403	.465	.181	.291
Dictation	.343	.152	−.455	.116	.684	−.559	.220	.231	.405	−.040	.128	−.553
Spanish Cloze 1	.014	−.711	−.171	.398	.256	.337	−.382	.413	.269	−.580	.383	.209
Spanish Cloze 2	−.019	−.458	.099	.284	.161	.202	−.435	.244	.106	−.384	.401	.339

angles as low as 4°, 6°, and 18°. These are well within the 95 percent Monte Carlo points for three common components and $n = 20$. This situation implies that the prime sources of variation among students in the two colleges are rather different, but most of the differences disappear if we characterize each college by three or more principal components. We see from the first two principal components of each college (Table 1) that the primary components for the two colleges load very differently on the variables structure and dictation. The compromise average subspace defined by the vectors in Table 2(b) presents a rather different pattern. Indeed, when component spaces of dimension four are compared, the vector nearest to both spaces loads almost exclusively on structure, with a secondary contribution from essay. This, then, is the component with respect to which the between-student scatter is most similar in the two colleges.

Now consider an overall comparison of all three colleges, using the technique of Section 3. For brevity, we will compare only the respective component spaces of dimension four. The results are given in Table 3.

Part (a) of Table 3 presents the coefficients of each of the four orthogonal directions closest to all three principal component subspaces, in relation to the original variables. Part (b) gives the angular separation between each group and each direction. The eigenvalues are the sum of squared cosines of these angles and are given at the foot of each column.

All three groups are close together along the first two axes, but begin to diverge after this. Inspection of the vector loadings gives an indication of the component with respect to which the between-student scatter is most similar in all three colleges. The first direction is loaded most heavily on structure (as in the comparison of colleges A and B) and on comprehension. This latter item was absent in the previous comparison, and it reflects the higher weighting given to this variable in the first two principal components of college C. The fact that it comes out as important when all colleges are considered together implies that it was hidden but not negligible in colleges A and B. The second direction in the present comparison is weighted almost exclusively on the Spanish Cloze tests (i.e., selection of a word from a list to complete a sentence) and corresponds to the third-most-important vector in Table 2. This indicates that only two dimensions are really common between the first four principal components of the three colleges. Structure and

2. Comparison of Principal Component Spaces of Colleges A and B

(a) Eigenvalue Analysis

	Eigenvalues				*Corresponding Angles in Degrees*			
Component spaces of dimension 1	.689				33.9			
Component spaces of dimension 2	.024	.739			81.1	30.7		
Component spaces of dimension 3	.173	.764	.977		65.5	29.1	8.7	
Component spaces of dimension 4	.521	.910	.989	.996	43.9	17.5	6.0	3.6

(b) Vectors Closest to Both Spaces

Dimension of comparison	*1*	*2*		*3*			*4*			
Vector	*1*	*1*	*2*	*1*	*2*	*3*	*1*	*2*	*3*	*4*
Comprehension	.227	.217	−.076	−.325	.194	.776	.181	−.753	.268	−.558
Essay	.365	.402	−.217	.382	.254	.288	.471	.075	.151	−.057
Cloze 1	.297	.280	.143	−.155	.394	−.217	.049	−.276	.030	.401
Cloze 2	.240	.242	−.006	−.085	.308	−.078	.092	−.210	.030	.248
Structure	.596	.599	.034	.318	.474	.310	.797	.094	−.468	−.016
Dictation	.537	.458	.586	−.278	.637	−.378	.077	−.449	−.013	.656
Spanish Cloze 1	.141	.254	−.644	.557	.130	−.146	.226	.193	.694	.180
Spanish Cloze 2	.074	.146	−.410	.480	.016	−.044	.206	.244	.451	.034

3. *Comparison of Principal Component Spaces of Colleges A, B, and C*

Dimension	*1*	*2*	*3*	*4*
(a) Directions Closest to Each Subspace				
Comprehension	.586	−.138	.247	−.756
Essay	.334	.223	.075	.246
Cloze 1	.219	.052	−.497	−.047
Cloze 2	.188	.048	−.375	−.029
Structure	.587	−.172	.350	.589
Dictation	.336	.076	−.608	.108
Spanish Cloze 1	.069	.769	.059	−.077
Spanish Cloze 2	.016	.547	.223	−.006
(b) Angles Formed by Each Group With Each Direction				
College A	1.99	9.38	17.48	15.01
College B	3.06	11.18	27.73	10.74
College C	1.57	6.33	29.71	73.75
Sum of Squared Cosines	2.995	2.924	2.448	1.977

comprehension form the major common cause of differences between students, followed by their performance on the Spanish tests. Differences among the colleges are highlighted by the fourth component, on which C diverges markedly from A and B. The difference is attributable to the contrast between structure and comprehension.

5. DISCUSSION

We conclude by noting a few characteristics of the methods given in this article and their similarities to other established techniques. A certain pattern in the eigenvalues is always obtained when different dimensionalities are successively compared. Clearly, if θ is the minimum angle between two spaces of dimension k, then the minimum angle between spaces of dimension greater than k for the same groups must be less than or equal to θ. An analogous argument will hold for the second-lowest angle and so on. Thus, in all tables having the form of 2(a), the diagonal and all subdiagonals of the triangular array of eigenvalues will be nondecreasing when going from left to right.

The geometrical connection between correlations and cosines of angles implies that there are close affinities between the techniques given earlier and canonical correlation analysis. In the latter case it is the variables rather than the units of the sample that are partitioned into two groups, and successive linear combinations (orthogonal within each group of variables) are produced to partition the overall correlations existing between the two groups. The basic geometrical argument underlying this, and hence the formal algebraic solution, has much in common with PCA comparison, although the emphasis has passed from units to variables. Canonical correlation analysis can also be extended to allow more than two sets of variables in the partition, and this area has been reviewed by Kettenring (1971). Unfortunately, there is no longer a unique model underlying the analysis, and Kettenring discussed a number of possibilities. The method given in Section 3 corresponds to Kettenring's sum of squared correlations model.

Finally, the concept of critical angles between subspaces, and allied geometry, has been used to advantage in the context of analysis of variance by James and Wilkinson (1971). It enables their factorization of the residual operator for nonorthogonal analysis of variance to be interpreted geometrically.

REFERENCES

Anderson, T.W. (1958), *An Introduction to Multivariate Statistical Analysis*, New York: John Wiley & Sons.
Cliff, N. (1962), "Analytical Rotation to a Functional Relationship," *Psychometrika*, 27, 283–295.
——— (1966), "Orthogonal Rotation to Congruence," *Psychometrika*, 31, 33–42.
Evans, G.T. (1971), "Transformation of Factor Matrices to Achieve Congruence," *British Journal of Mathematical and Statistical Psychology*, 22, 24–48.
Gower, J.C. (1967), "Multivariate Analysis and Multidimensional Geometry," *The Statistician*, 17, 13–28.
Horst, P. (1956), "A Simple Method of Rotating a Factor Matrix to a Simple Structure Hypothesis," *Journal of Experimental Education*, 24, 251–258.
Hurley, J.R., and Cattell, R.B. (1962), "The Procrustes Program: Producing a Direct Rotation to Test a Hypothesized Factor Structure," *Behavioural Science*, 7, 258–262.
James, A.T., and Wilkinson, G.N. (1971), "Factorization of the Residual Operator and Canonical Decomposition of Nonorthogonal Factors in the Analysis of Variance," *Biometrika*, 58, 279–294.
Kendall, M.G., and Stuart, A.S. (1968), *The Advanced Theory of Statistics* (Vol. 3) (2nd ed.), London: Griffin.
Kettenring, J.R. (1971), "Canonical Analysis of Several Sets of Variables," *Biometrika*, 58, 433–451.
Korth, B.A. (1973), "Analytic and Experimental Examination of Factor Matching Methods," unpublished dissertation, University of Illinois at Urbana-Champaign.
———, and Tucker, L.R. (1976), "Procrustes Matching by Congruence Coefficients," *Psychometrika*, 41, 531–535.
Lawley, D.N., and Maxwell, A.E. (1971), *Factor Analysis As a Statistical Method* (2nd ed.), London: Butterworth.
Schonemann, P.H., and Carroll, R.M. (1970), "Fitting One Matrix to Another Under Choice of a Central Dilation and Rigid Motion," *Psychometrika*, 35, 245–255.
Seal, H. (1964), *Multivariate Statistical Analysis for Biologists*, London: Methuen and Co.