DANIEL F. MERRIAM, Editor

MULTIVARIATE PROCEDURES AND FORTRAN IV PROGRAM FOR EVALUATION AND IMPROVEMENT OF CLASSIFICATIONS

By

FERRUH DEMIRMEN

Stanford University

COMPUTER CONTRIBUTION 31

State Geological Survey

The University of Kansas, Lawrence

1969

in cooperation with the
American Association of Petroleum Geologists
Tulsa, Oklahoma

Editor's Remarks

With the publication of Computer Contribution 31, we begin the fourth year of the series. Since 1966 more than 60,000 copies of these publications have been distributed in 40 countries. The series now is sponsored jointly by the Geological Survey and the American Association of Petroleum Geologists, the largest geological organization in the world.

In 1968, about 125 computer programs were made available to workers. The use of the new techniques is becoming widespread and routine and very successful in many instances. Todate most geological applications have been in statistics, trend analysis, classification and more recently simulation. Undoubtedly other applications will be found.

This program "Multivariate procedures and FORTRAN IV program for evaluation and improvement of classifications" by Ferruh Demirmen lists criteria by which different classifications can be judged as to their efficiency. For a limited time the program described here will be made available on magnetic tape for $15.00. An extra $10.00 is charged if punched cards are required.

For an up-to-date list of COMPUTER CONTRIBUTIONS write the Editor, Kansas Geological Survey, The University of Kansas, Lawrence, Kansas, 66044, U.S.A.

---

* Active Member, ° Associate Member, + Junior Member, American Association of Petroleum Geologists.

MULTIVARIATE PROCEDURES AND FORTRAN IV PROGRAM FOR EVALUATION AND IMPROVEMENT OF CLASSIFICATIONS

by

Ferruh Demirmen

ABSTRACT

ITERIM is an IBM System/360 FORTRAN IV(H) program designed primarily to assess and improve classifications, although it can be used also for principal component analysis, discriminant analysis, and one-way multivariate analysis of variance. Three criteria, pooled within-groups sum of squares, Wilks' Lambda, and the sum of the eigenvalues associated with discriminant functions, are computed to assess and compare classifications. The improvement of a classification is achieved through reduction of the pooled within-groups sum of squares in the discriminant space. The classifications compared must contain the same number of items, the same number of groups, and must be defined relative to the same number of variables. A number of options, both as to computations and output, are provided.

INTRODUCTION

Geologists and others dealing with multivariate classification or "cluster analysis" are faced frequently with a great diversity of techniques from which to choose (Sokal and Sneath, 1963; Ball, 1965; Williams and Dale, 1965; Fortier and Solomon, 1966; Goodall, 1966a, 1966b; Gower, 1967a; Johnson, 1967). Some of these techniques concern weighting or standardization of data, others concern similarity measures, and yet others are related to grouping of data. At present there exists little a priori rational basis for choosing between these diverse techniques, although a number of writers (Sokal and Rohlf, 1962; Eades, 1965; Minkoff, 1965; Rohlf and Sokal, 1965; Gower, 1967b) have discussed the merits and demerits of certain techniques. With different clustering techniques, the resulting classifications will be different, and it may be difficult to reconcile the conflicting classifications. A way out of this dilemma seems to be the use of a variety of techniques and evaluate, in retrospect, the resulting classifications. Such evaluation can be made either on a substantive and subjective basis, or alternatively, on an objective basis. Furthermore, it would be desirable if any of the classifications obtained by cluster analysis could be further improved by some criterion.

The computer program (ITERIM) presented here is designed primarily to evaluate and improve classifications by objective criteria.* In addition, as intermediate steps, the program computes principal components and multiple linear discriminant functions and performs a one-way multivariate analysis of variance. Techniques used for evaluation and improvement are nonprobabilistic in nature. It is assumed that data on which a classification is based are metric in nature, that is they consist of measurements taken on a continuous scale. For nonmetric or semiquantitative data other techniques of evaluation and improvement might be more appropriate, although, as an exploratory tool, the program may be useful for such data as well. The program accepts a classification as input. It does not do cluster analysis; nor does it assign a new item to a class. In computing the principal components, the classes are ignored and the data are treated as a whole. A number of options, both as to computations and output, are provided.

The ITERIM program described here is an outgrowth of the program originally given by Casetti (1964). The criteria used for evaluation and improvement of a classification are the same as those which Friedman and Rubin (1967) employ to "optimize" a partition in cluster analysis, although the ITERIM was written before Friedman and Rubin's paper was published. The papers of Forgy (1965) and MacQueen (1966) also are cognate with the techniques utilized in the program.

The writer is indebted to Dr. Paul Switzer for many valuable and stimulating discussions, and to Drs. J.E. Klovan and F.J. Rohlf for helpful editorial suggestions. All statements herein, however, are the responsibility of the writer. Partial financial support for the development of the program was provided by a NATO Science Fellowship to the writer and by a National Science Foundation grant (NSF GP 4514) to Dr. J.W. Harbaugh. The School of Earth Sciences of Stanford University furnished most of the computer time.

---

*It is recognized that the word "objective" is a relative term, and the selection of a so-called objective criterion for the evaluation or improvement of a classification involves a certain amount of subjective judgment on the part of the investigator.

MATHEMATICAL DEVELOPMENT

Preliminaries

In the text that follows a small letter with a bar sign ("‑") underneath will denote a vector, a capital letter with the same sign below will denote a matrix, and a letter without this sign will denote a scalar. If β_i $(i = 1, \ldots, p)$ are a set of scalars, then $\underline{D}(\beta_i)$ will designate a $(p \times p)$ diagonal matrix whose principal diagonal elements are the scalars β_i arranged in descending order according to i, that is $d_{ij} = 0$ $(i \neq j)$ and $d_{ii} = \beta_i$. Furthermore, if some matrix $\underline{\Omega}(n \times p)$ contains the scores of n items with respect to p variables, these variables will be referred to as the ω-variables, and ordinary Euclidean space identified by them will be referred to informally as the ω-space. The i-th row vector of $\underline{\Omega}$ then represents the i-th item and can be thought of as a point in the p-dimensional Euclidean ω-space. The variables (space) which form the basis of evaluation and improvement of a classification will be designated as <u>initial</u> variables (space), which may or may not be <u>identical</u> to <u>input</u> variables (space). All correlations and discriminant functions will be understood to be product-moment correlations and linear discriminant functions, respectively.

Evaluation of a Classification

Evaluation of a classification is made on the basis of three criteria that purport to measure the quality of a classification. The three criteria measure, in three different senses,* the degree of "compactness" of a classification, so that the quality of a classification is equated with its "compactness." Of any two classifications, the one that is more "compact" by a given criterion is regarded "better" relative to that criterion. The meaning of "compactness" will be evident in discussion of the criteria. The three criteria are not related monotonically, so that a classification which ranks "best" among a number of classifications by a particular criterion need not rank as the "best" by the other two criteria, although in general it might be expected. The decision to choose among the three criteria is left to the investigator and introduces an element of subjectivity into the evaluation process. As will be noted below, however, two of the criteria (tr $\underline{W}$ and Λ) have, in the writer's own experience, given consistent rankings and may be recommended tentatively in preference to the third one (tr $\underline{W}^{-1}\underline{B}$). For a given classification the program computes all three criteria. The three criteria have been used by Friedman and Rubin (1967) to "optimize" a partition in cluster analysis.

To use of our evaluation criteria requires that the two classifications that are being compared contain the same number of items and the same number of groups, and be defined relative to the same number of variables. The initial scores in the two classifications need not be the same, provided that cognizance is made of the problem of invariancy of the criteria. In the discussion that follows it is assumed that the classifications compared meet the requirement noted above.

*Analogous to the way that the median, the arithmetic mean, and the geometric mean measure, in three different senses, the "central tendency" of a variable.

Scatter Matrices

The three criteria of evaluation as defined are based on the within, between, and total scatter matrices (in the sense of Wilks, 1960, 1962). Assume that a classification represents the partition of n items into m groups on the basis of p variables, with the h-th group containing n_h items. Hence $\Sigma n_h = n$. Let the initial score matrix that identifies this classification and serves as the basis of classificatory analysis be the partitioned matrix $\underline{X} = (x_{hki})$ $(h = 1, \ldots, m; k = 1, \ldots, n_h; i = 1, \ldots, p)$ whose element x_{hki} is the score of the k-th item on the i-th variable, with the k-th item being contained in the h-th group. Let

$$x_{h.i} = \frac{1}{n_h} \sum_{k=1}^{n_h} x_{hki}$$

be the mean of the i-th variable (x_i) over the h-th group, and

$$x_{..i} = \frac{1}{n} \sum_{h=1}^{m} n_h x_{h.i} = \frac{1}{n} \sum_{h=1}^{m} \sum_{k=1}^{n_h} x_{hki}$$

be the grand mean of the i-th variable of the n items. Then the matrices $\underline{W} = (w_{ij})$, $\underline{B} = (b_{ij})$, $\underline{T} = (t_{ij})$, $i, j = 1, \ldots, p$, where

$$w_{ij} = \sum_{h=1}^{m} \sum_{k=1}^{n_h} (x_{hki} - x_{h.i})(x_{hkj} - x_{h.j}),$$

$$b_{ij} = \sum_{h=1}^{m} n_h (x_{h.i} - x_{..i})(x_{h.j} - x_{..j}), \text{ and}$$

$$t_{ij} = \sum_{h=1}^{m} \sum_{k=1}^{n_h} (x_{hki} - x_{..i})(x_{hkj} - x_{..j}),$$

represent, respectively, the within, between, and total scatter matrices in the x-space. A more cumbersome name for $\underline{W}$ is the "within-groups sum of squares and cross products matrix"; and similarly for $\underline{B}$ and $\underline{T}$. Note that $\underline{T} = \underline{W} + \underline{B}$.

Trace W Criterion

$$\text{tr } \underline{W} = \sum_{h=1}^{m} \sum_{k=1}^{n_h} \sum_{i=1}^{p} (x_{hki} - x_{h.i})^2 \qquad (1)$$

is the total within-groups sum of squares with respect to all p variables pooled over all m groups, hence a reasonable criterion to assess the quality of a classification. A classification associated with a small tr $\underline{W}$ value can be regarded "compact" in the sense that total variability within groups about the respective means is small. Because tr $\underline{W}$ + tr $\underline{B}$ = tr $\underline{T}$ = constant in the x-space, small tr $\underline{W}$ is equivalent to a large tr $\underline{B}$ or a large tr $\underline{B}$/tr $\underline{W}$. Trace $\underline{B}$ in effect represents the weighted sum (weighted by group sizes) of squared ordinary Euclidean distances between group centers of gravity and the grand center of gravity. Thus small tr $\underline{W}$ also implies that the total variability among the groups is large, that is centers of gravity of the groups are dispersed from the grand center of gravity. It follows that, of two classifications, the one having a small tr $\underline{W}$ value, or equivalently, a large tr $\underline{B}$ or a large tr $\underline{B}$/tr $\underline{W}$ value, can be regarded "better." Note should be made that tr $\underline{W}$ does not take into account group covariances. In general group covariances will be nonzero if the total covariances (measured over all n items) are zero, so that transforming initial variables into a set of uncorrelated variables does not help. It is easy to show that tr $\underline{W}$, tr $\underline{B}$ and tr $\underline{T}$ are invariant under orthogonal transformations.

Wilks' Lambda Criterion

The determinantal ratio,

$$\Lambda = \frac{|\underline{W}|}{|\underline{T}|}, \qquad (2)$$

is a scalar quantity that was proposed by Wilks (1932) as a statistic to test equality of group mean vectors under assumption of normality and equal group covariance matrices. Λ represents the ratio of within-groups scatter to total scatter (Wilks, 1960, 1962), and be regarded as another measure of the quality or "compactness" of a classification, with small Λ values corresponding to a "good" classification. Except for the degrees of freedom, Λ also represents the ratio of within to total generalized variance in the sample. For a geometric interpretation of generalized variance, see Anderson (1958). Λ is invariant under all nonsingular linear transformations and in this respect has advantage over the tr $\underline{W}$ criterion. The use of the Λ criterion, however, requires the nonsingularity of $\underline{W}$, which in turn requires that $p \leq n-m$ (assuming the p variables are linearly independent). If the number of variables is too large to meet this requirement, then orthonormalization (see below) can be used to reduce the number of variables before performing classificatory analysis. Note that when $\underline{W}$ is nonsingular (positive definite), so is $\underline{T}$. It is easy to see that $\Lambda = 1/|\,\underline{W}^{-1}\underline{B} + \underline{I}\,|$.

It may be added that the F-statistic, also used to test the equality of group mean vectors and computed in the program, is a decreasing monotonic function of Λ, so that two classifications can be compared also on the basis of their F-values. In this situation the "better" classification will be associated with the larger F-value. The ratings of the classifications would of course be the same as with the Λ criterion.

Trace $W^{-1}B$ Criterion

$$\text{tr } \underline{W}^{-1}\underline{B} = \sum_{h=1}^{m} n_h \sum_{i,j=1}^{p} w^{ij} (x_{h.i} - x_{..i})(x_{h.j} - x_{..j}), \qquad (3)$$

where w^{ij} is (i, j)-th element of $\underline{W}^{-1}$, represents the weighted sum (weighted by group sizes) of squared Mahalanobis distances between group centers of gravity and the grand center of gravity, and is equivalent to what Rao (1952, p. 257) has called generalization of the Mahalanobis D^2 to more than two groups. The trace of $\underline{W}^{-1}\underline{B}$ has been used also as a test statistic in the instance of the general linear hypothesis under the assumption of normality and a common covariance matrix, with larger values of tr $\underline{W}^{-1}\underline{B}$ leading to an easier rejection of the null hypothesis (Hotelling, 1951; Anderson, 1958). It is reasonable to regard tr $\underline{W}^{-1}\underline{B}$ therefore as another measure of the "compactness" of a classification, with larger values of tr $\underline{W}^{-1}\underline{B}$ indicating a more "compact" or "better" classification, whereby group mean vectors are dispersed about the grand mean vector. Unlike tr $\underline{B}$, to which it is analogous, tr $\underline{W}^{-1}\underline{B}$ has the intuitively appealing property that it corrects for correlations between groups. In working with actual data, however, the writer found that ratings of classifications by the tr $\underline{W}^{-1}\underline{B}$ criterion were somewhat erratic relative to ratings by the Λ and tr $\underline{W}$ criteria, which were by and large in agreement. If this can be taken as a tentative indication of the relative merits of our evaluation criteria, it follows that use of the tr $\underline{W}^{-1}\underline{B}$ criterion might be discouraged. For purposes of cluster analysis, Friedman and Rubin (1967) also favor the Λ criterion over the tr $\underline{W}^{-1}\underline{B}$ criterion, although they are ambivalent about the tr $\underline{W}$ criterion. Like the Λ criterion, tr $\underline{W}^{-1}\underline{B}$ criterion is invariant under all nonsingular linear transformations, and its use requires that $p \leq n-m$.

The ITERIM program takes advantage of the symmetricity of $\underline{W}$ and $\underline{B}$ and computes tr $\underline{W}^{-1}\underline{B}$ by a special procedure which does not require the inversion of $\underline{W}$.

Discriminant Functions

The improvement of a classification in the program is performed in the discriminant space, so that a brief discussion of these functions is germane at this point. Discriminant functions are useful for concentrating the total discriminatory power of x-variables in $\tilde{p}$ dimensions, where $\tilde{p} \leq p$, or for obtaining a new set of orthogonal coordinate axes along which variation between groups is maximized relative to variation within groups.

Let γ_i ($i = 1, \ldots, p$) be the i-th eigenvalue of $\underline{W}$, and $\underline{M}$ ($p \times p$) an orthogonal matrix whose columns are normalized eigenvectors of $\underline{W}$ arranged in the same order as γ_i. Assuming that $\underline{W}$ is nonsingular, let $\underline{K}$ ($p \times p$) be a symmetric matrix such that

$$\underline{K} = \underline{D}^{-1}(\sqrt{\gamma_i})\,\underline{M}'\;\underline{B}\,\underline{M}\,\underline{D}^{-1}(\sqrt{\gamma_i}).$$

We recall that matrices $\underline{W}$ and $\underline{B}$ are both defined in the initial x-space. Let θ_i ($i = 1, \ldots, p$) be the i-th eigenvalue of $\underline{K}$, and $\underline{R}$ ($p \times p$) an orthogonal matrix whose columns contain normalized eigenvectors of $\underline{K}$ in the same order as θ_i. Then it can be shown, from similarity relations of matrices, that θ_i's are also eigenvalues of $\underline{W}^{-1}\underline{B}$, and that nonsingular matrix $\underline{V}$ ($p \times p$), where

$$\underline{V} = \underline{M}\,\underline{D}^{-1}(\sqrt{\gamma_i})\,\underline{R},$$

contains, in its columns, a set of eigenvectors of $\underline{W}^{-1}\underline{B}$. Furthermore, matrix $\underline{V}$ simultaneously diagonalizes $\underline{W}$ and $\underline{B}$ such that

$$\underline{V}'\,\underline{W}\,\underline{V} = \underline{I}, \quad \text{and} \quad \underline{V}'\,\underline{B}\,\underline{V} = \underline{D}(\theta_i). \qquad (4)$$

Thus, if we denote the z-th column of $\underline{V}$ as $\underline{v}_z$, it follows that

$$\underline{v}_z'\,\underline{B}\,\underline{v}_z = \theta_z, \quad \underline{v}_z'\,\underline{W}\,\underline{v}_z = 1, \text{ and}$$

$$\theta_z = \frac{\underline{v}_z'\,\underline{B}\,\underline{v}_z}{\underline{v}_z'\,\underline{W}\,\underline{v}_z} = \frac{\sum_{i,j=1}^{p} v_{iz}\, v_{jz}\, b_{ij}}{\sum_{i,j=1}^{p} v_{iz}\, v_{jz}\, w_{ij}}. \qquad (5)$$

Clearly, $\sum_{z=1}^{p} \theta_z = \text{tr } \underline{W}^{-1}\underline{B}$, which is the way this criterion is computed in the program.

Discriminant functions are obtained by the transformation $\underline{Y} = \underline{X}\,\underline{V}$, where $\underline{Y}$ ($n \times p$) contains scores of n items with respect to p discriminant functions (y-variables). If we express b_{ij} and w_{ij} of (5) in terms of x-variables (see "Scatter Matrices"), and note that y-variables are linear combinations of x-variables with elements of $\underline{v}_z$ ($z = 1, \ldots, p$) as the coefficients, it is easy to see that (5) is equivalent to

$$\theta_z = \frac{\sum_{h=1}^{m} n_h\,(y_{h.z} - y_{..z})^2}{\sum_{h=1}^{m}\sum_{k=1}^{n_h} (y_{hkz} - y_{h.z})^2}. \qquad (6)$$

Thus, θ_z is the ratio of between- to within-groups sum of squares in the z-th discriminant dimension. In Wilks' (1960) terminology θ_z is the ratio of between to within scatter in the same dimension. Hence, θ_z can be regarded as the discriminatory power of the z-th discriminant function.

We can, without loss of generality, arrange discriminant functions in order of relative magnitudes of associated eigenvalues, so that the first discriminant function has the greatest discriminatory power, the second the next highest discriminatory power, etc. The number of nonzero eigenvalues of $\underline{W}^{-1}\underline{B}$ is equal to the rank of $\underline{W}^{-1}\underline{B}$, which is also the rank of $\underline{B}$. Let $\tilde{p}$ be this number. Then, assuming that p variables in the x-space are linearly independent, $\tilde{p}$ is the lesser of $(m-1)$ and p. Hence, if $p > m-1$, the total discriminatory power of initial variables will be contained in fewer than p discriminant dimensions, which provides a nice parsimony in dimensionality. A measure of the cumulative power of the first, say z ($\leq p$), discriminant functions, is given by

$$\sum_{i=1}^{z} \theta_i \Big/ \sum_{i=1}^{p} \theta_i = \sum_{i=1}^{z} \theta_i \Big/ \text{tr } \underline{W}^{-1}\underline{B}.$$

An alternate, and probably more meaningful, measure of the cumulative power associated with the first z discriminant functions is

$$\prod_{i=1}^{z} (1 + \theta_i)^{-1}. \qquad (7)$$

This expression represents the ratio of within-groups scatter to total scatter in the z-dimensional discriminant space. When $z = \tilde{p}$ (and of course, also when $z = p$), (7) gives the ratio of within-groups scatter to total scatter in the p-dimensional initial space. This ratio, it will be recalled, is our familiar Wilks' Lambda, Λ. The program takes advantage of this fact and computes Λ from the formula (7), setting $z = p$. We note, in passing, that discriminant functions are uncorrelated, with the z-th discriminant function having a sample variance $(1 + \theta_z) / (n-1)$.

In our discussion, it was assumed that column vectors of $\underline{V}$ are left nonnormalized, which is the usual procedure of computing discriminant functions. The ITERIM program allows an option to normalize these vectors. If we let $\underline{Y}^*$ be the counterpart of $\underline{Y}$ when these eigenvectors are normalized, then $\underline{Y}^*$ represents a diagonal transformation of $\underline{Y}$,

$$Y^* = \underline{Y}\, \underline{D}^{-1}(l_i) \quad ,$$

where l_i is the length of the i-th column of $\underline{V}$. With normalization the relations in (4) do not hold (unless $\underline{W} = \underline{I}$), although eigenvalues, of course are not affected. Furthermore, θ_z represents the ratio of between- to within-groups sum of squares in the z-th discriminant dimension. Normalization, although altering variances of discriminant functions, does not affect their uncorrelatedness. The importance of normalization in connection with improvement of a classification will be noted under "Discussion."

It is useful to compute correlations between input variables and discriminant functions. These correlations give a measure of the "importance" or "weight" of each input variable on each discriminant function. The program computes these correlations by the usual formula (given here), but takes advantage of the fact that, if the eigenvectors of $\underline{W}^{-1}\underline{B}$ are nonnormalized the variance of a discriminant function is a simple function of the associated eigenvalue, and if these vectors are normalized and the data are orthonormalized, variances are uniformly $1/(n-1)$ (see Appendix). It can be verified that these correlations remain invariant with scale alteration of input data, with orthonomalization (see below) if all principal components are retained, and with normalization of eigenvectors of $\underline{W}^{-1}\underline{B}$.

Improvement of a Classification

The rationale behind improvement of a classification is a logical extension of the concept of evaluation. We can alter a classification in such a way that the altered classification will rate "better" by a particular criterion. Hence, by this principle, a rearrangement of a classification so as to reduce tr $\underline{W}$ in a given space marks an improvement in that space relative to that criterion. Similarly, a rearrangement resulting in a reduction of Λ or an increase in tr $\underline{W}^{-1}\underline{B}$ represents an improvement relative to these criteria. Inasmuch as the three criteria are not related monotonically, an improvement relative to a given criterion need not mark an improvement relative to the other two criteria, although in general it would be expected that this be the situation. The criterion which the ITERIM program utilizes to improve a classification is the tr $\underline{W}$ criterion. From the computational point of view, improvement by this criterion is the easiest to perform.

Nearest Neighbor Algorithm

An efficient method of improving a classification by the tr $\underline{W}$ criterion is provided by the nearest neighbor algorithm, whereby each item is allocated to that group to which it is nearest in terms of ordinary Euclidean distances. The procedure is analogous to the "k-means" method of MacQueen (1966). Each group is represented by its center of gravity, that is the mean vector computed for that group. Although the algorithm can be designed to operate in any arbitrary space, the ITERIM program allows the algorithm to operate in the discriminant space generated from the initial space. The reason for this will be evident under "Discussion." In computing distances, all discriminant functions are used. Hence, returning to our notation, if we let $\underline{y}_{hk.}$ ($1 \times p$) be the vector representing the k-th item in the h-th group in the discriminant space, and

$$\underline{y}_{h..} = \frac{1}{n_h} \sum_{k=1}^{n_h} \underline{y}_{hk.}$$

the mean vector for the h-th group in the same space, then the nearest neighbor algorithm assigns the item in question to the g-th group for which the distance

$$[(\underline{y}_{hk.} - \underline{y}_{g..})(\underline{y}_{hk.} - \underline{y}_{g..})']^{1/2}$$

is smallest for all $g = 1, \ldots, m$. If $g = h$, the item remains in its group; otherwise it is displaced to the g-th group. This procedure is repreated for all n items. The displacement of items from their original groups creates a new classification, whereupon new mean vectors are recomputed. These steps are repeated iteratively, with each iteration yielding a new classification generated from the partition of the immediately preceding iteration. If we let $\underline{W}_{(y)}$ be the within scatter matrix in the discriminant space (y-space), it is evident that this method of reshuffling items during each iteration reduces tr $\underline{W}_{(y)}$, thus marking an improvement in classification relative to the tr $\underline{W}$ criterion in the discriminant space. Hence the iterations produce successive improvements in classification by means of incremental reduction in tr $\underline{W}_{(y)}$.

Iterations are terminated when an improvement by the nearest neighbor algorithm is no longer possible, or when the maximum number of iterations specified by the user is exceeded. If the improvement is no longer possible, the final classification can be considered a "stabilized" form of input classification. When a "stabilized" condition occurs, partitions obtained during the last two iterations are, of necessity, identical. The classifications obtained during iterations are influenced by the arrangement of items in the input classification (for details, see Casetti, 1964). The number of groups (m) remains unchanged during the iterations.

Core Items

A measure of the "distance" between input classification and classification generated during an iteration is given by the number of "core items." A core item is that item which, at the end of a given iteration, is found in the same group as it was in the input classification. Hence, a large number of core items, indicating relatively little reshuffling of items from their original groups, suggests that the classification obtained during the current iteration is not too "distant" from the input classification. The user, however, is cautioned against attaching much significance to the concept of core items. It is certainly more meaningful to compare the current classification and the input classification by the criteria which we gave earlier, than by the relative number of core items. Ordinarily the number of core items decreases as iterations proceed, although slight reversals may occur.

One-way Analysis of Variance

To test the null hypothesis that group populations have equal mean vectors, the program uses the F-approximation given by Rao (1952, p. 258-262). The test assumes that group populations are normally distributed with a common covariance matrix. Then, with these assumptions, the statistic

$$F = \frac{1 - \Lambda^{1/s}}{\Lambda^{1/s}} \cdot \frac{ks + 2\lambda}{2r}, \quad \text{where}$$

$$s = \frac{p^2(m-1)^2 - 4}{p^2 + (m-1)^2 - 5}, \quad k = n - 1 - \frac{p+m}{2},$$

$$\lambda = -\frac{p(m-1) - 2}{4}, \quad r = \frac{p(m-1)}{2},$$

can be used as a variance ratio with $(2r)$ and $(ks + 2\lambda)$ degrees of freedom. Quantities n, p, m, and Λ are the same as we have been using throughout our discussion. Note that, since Λ is invariant under nonsingular linear transformations, so is F, although the program computes F in the x-space. When $p = 1$, Λ becomes a mere ratio of within-groups to total sum of squares, and the F-statistic is reduced to its familiar form in the univariate case:

$$F = \frac{B_{ss}}{W_{ss}} \cdot \frac{n-m}{m-1},$$

where B_{ss} and W_{ss} are the between- and within-groups sum of squares, respectively, Before making decisions on the basis of the F-values, it is well to check assumptions of normality and equal covariance, which the present program does not do. To test the homogeneity of covariances, the program given by Wolleben, Pauken, and Dearien (1968) may be used.

Optional Transformations Prior to Classificatory Analysis

Up to this point it was assumed tacitly that the initial score matrix $\underline{X}$ (n×p) serving as the basis of evaluation and improvement of a classification is an input data matrix. It may be desirable in some instances to transform input data before the classificatory analysis is performed. Two such transformations, scale alteration and orthonormalization, provided as options in the program, are described. Either one or both of the transformations can be performed on the input data. In either instance, the input score matrix will be designated as some matrix other than $\underline{X}$, as noted.

Scale Alteration

The option of scale alteration is provided chiefly to enable the user to suppress the scales of his input variables so that results will be printed or punched in fields specified by output formats. Furthermore, output formats are designed in such a way that, if the number of variables is 14 or less, and the number of groups 23 or less, results will be printed in easy-to-read tables (for example, the within scatter matrix will not be separated). These features place a constrain on the scales (variances) of input variables. As a rule-of-thumb, input variables should not have variances greatly in excess of $1/n$ (n = total number of items). If the input scores do not meet this requirement, then their scales (variances) should be readjusted. However, if the data are orthonormalized (see below), such scale readjustment will in most instances be unnecessary. In addition to the obtainment of a readable output, the user also may wish to alter the scales of his input variables for reasons of his own before classificatory analysis is performed, so that scale alteration is a useful and convenient option.

To formulate, let the input data contain n items each characterized by its measurements with respect to q ($\geq$ p) variables, which we designate as the $e^{(1)}$-variables. Disregarding partitioning of data into groups, let $\underline{E}^{(1)}$ (n q) be the score matrix identifying input data. If c_i (i= 1, . . . , q) are some positive constants supplied by the user, then the diagonal transformation

$$\underline{E}^{(2)} = \underline{E}^{(1)} \underline{D}^{-1} (\sqrt{c_i}) \tag{8}$$

alters variances of $e^{(1)}$-variables and yields a set of new $e^{(2)}$-variables such that

$$\text{var}(e_i^{(2)}) = \frac{1}{c_i} \text{var}(e_i^{(1)})$$

for all $i = 1, \ldots, q$. The matrix $E^{(2)}$ $(n \times q)$ is the new score matrix with respect to $e^{(2)}$-variables. Alteration of variances can be regarded as a change in the scales of input variables. If any $c_i > 1$, then respective scale alteration will mean reduction in scale (variance) of $e_i^{(1)}$. If c_i's are all equal, say to some constant c (> 0), then scale alteration will be uniform for all input variables, and the scales (variances) will be altered by a constant factor of $1/c$. Unless there are special reasons to do otherwise, scale alteration should be uniform for all input variables (see "Discussion").

Now, let the matrix $\underline{E}$ $(n \times q)$ stand for $\underline{E}^{(2)}$ if scale alteration is requested, and for $\underline{E}^{(1)}$ if this option is bypassed. If orthonormalization, described below, is requested, it is based on $\underline{E}$; if not, the program assumes that $\underline{E}$ is the initial score matrix for purposes of classificatory analysis, that means in our notation, sets $p = q$ and $\underline{X} = \underline{E}$.

Orthonormalization

Orthonomalization, provided as an option in the program, refers to a series of linear transformations whereby e-variables are transformed to a new set of uncorrelated variables each with mean zero and variance $1/(n-1)$. The orthonormalized data can then be used for classificatory analysis, that is, to evaluate and improve a classification. Furthermore, the number of variables can be reduced before performing classificatory analysis. Thus, orthonomalization has the dual purpose of (i) obtaining uncorrelated variables with equal variance, and (ii) reducing the number of input variables.

Orthonomalization is based on the $\underline{E}$ $(n \times q)$ score matrix and is achieved through principal components generated from the e-space. Principal components can be extracted either from the covariance matrix or correlation matrix, and both options are provided. The decision between covariance and correlation options is left to the user and should be made chiefly on substantive grounds. If e-variables are all measured in the same or comparable units (for example, all measuring weight in grams), then standard procedure in component analysis is to use the covariance option. This is because, as we shall shortly see, principal components are linear combinations of "original" e-variables that tend to contain large portions of the total variance, and if the correlation option is used, the total "variance," being the number of e-variables (= q), has a rather artificial quality. Furthermore, as Anderson (1963) has shown, the sampling theory of principal components under the correlation option is much more complicated than its counterpart under the covariance option. If e-variables are measured in noncomparable units, however, the rationale behind the covariance option becomes highly dubious, and in such cases the usual recourse is to use the correlation option.

It must be stressed at this point that the foregoing remarks about covariance and correlation options are germane only insofar as the chief interest is in principal components per se, that is when ITERIM is used primarily for principal component analysis. As we shall see under "Discussion," provided that all principal components are retained, either option leads to the same conclusion as far as evaluation and improvement of a classification, so that under these circumstances the question of choosing between covariance and correlation options becomes purely academic.

To formulate our approach, let $e_{\cdot i} = n^{-1} \sum_{k=1}^{n} e_{ki}$ be the mean of the variable e_i over the n items, and $\overline{\underline{E}}$ a $(n \times q)$ matrix whose i-th column contains uniformly the mean $e_{\cdot i}$. Then the matrix $\underline{F}$ $(n \times q)$, where

$$\underline{F} = \underline{E} - \overline{\underline{E}} = (e_{ki} - e_{\cdot i}) \,, \quad k = 1, \ldots, n; \quad i = 1, \ldots, q \,,$$

represents the matrix of deviations from the means and defines a new set of f-variables such that, for each variable f_i, the mean $f_{\cdot i} = 0$. If we let s_i, assumed to be positive, be the standard deviation of e_i (also of f_i), then the matrices $\underline{C}$ $(q \times q)$ and $\underline{R}$ $(q \times q)$, where

$$\underline{C} = \frac{1}{n-1} \underline{F}' \underline{F} \quad \text{and} \quad \underline{R} = \underline{D}^{-1}(s_i) \, \underline{C} \, \underline{D}^{-1}(s_i) \,,$$

represent, respectively, covariance and correlation matrices in the e-space (also in the f-space). The transformation

$$\underline{Z} = \underline{F} \, \underline{D}^{-1}(s_i)$$

standardizes f-variables and yields a new set of z-variables each with variance 1 (and mean 0). The matrix $\underline{Z}$ $(n \times q)$ is the score matrix with respect to z-variables. Note that $\underline{R} = (n-1)^{-1} \underline{Z}' \underline{Z}$, that is $\underline{R}$ is the covariance as well as the correlation matrix in the z-space. Note, also, that $\underline{C}$ and $\underline{R}$ are both symmetric.

(1) Covariance option: Let λ_i be the i-th ($i = 1, \ldots, q$) largest eigenvalue of $\underline{C}$, and $\underline{A}$ (qxq) an orthogonal matrix whose i-th column is the normalized i-th eigenvector of $\underline{C}$. We shall assume, for simplicity, that λ_i's are distinct. Then $\underline{A}$ is uniquely determined. The transformation $\underline{G} = \underline{F}\ \underline{A}$ linearly maps the f-space into a g-space defined by principal components. The matrix $\underline{G}$ (nxq) is the score matrix with respect to these principal components. It can be readily shown that principal components are uncorrelated, with the i-th principal component g_i having the variance λ_i, and the sum of variances of q principal components equaling tr $\underline{C}$, the total variance contained in the f-space (also in the e-space). Thus the i-th principal component is that linear compound of "original" e-variables which contains the i-th largest portion of the total variance. The proportion of total variance attributable to i-th principal component is $\lambda_i / \text{tr } \underline{C}$, and the proportion of cumulative variance associated with the same component is $\sum_{j=1}^{i} \lambda_j / \text{tr } \underline{C}$.

In some applications of principal components, it may be desirable to compute correlations between input variables ($e^{(1)}$-variables) and principal components. These correlations give a measure of "importance" of each input variable on each principal component. Computation of these correlations is provided as an option in the program. If we let $\underline{R}^*$ (qxq) be the matrix whose element r^*_{ij} is the correlation between the variable f_i and the principal component g_i, we have

$$\underline{R}^* = \frac{1}{n-1}\ \underline{D}^{-1}(s_i)\ \underline{F}'\ \underline{G}\ \underline{D}^{-1}(\sqrt{\lambda_i}).$$

Noting that, from the theory of symmetric matrices, $\underline{A}'\ \underline{C}\ \underline{A} = \underline{D}(\lambda_i)$, and recalling that $\underline{C} = (n-1)^{-1}\underline{F}'\ \underline{F}$ and $\underline{G} = \underline{F}\ \underline{A}$, we obtain

$$\underline{R}^* = \underline{D}^{-1}(s_i)\ \underline{A}\ \underline{D}(\sqrt{\lambda_i})\ .$$

This equation provides a convenient method of computing correlations between f-variables and principal components. Inasmuch as f-variables differ from $e^{(1)}$- and $e^{(2)}$-variables only in origin, and the last-named two differ from each other only in scale, it follows that element r^*_{ij} of $\underline{R}^*$ represents not only correlation between f_i and g_i, but also between $e^{(2)}_i$ and g_i, and between $e^{(1)}_i$ and g_i. Note that, unlike $\underline{C}$ and $\underline{R}$, $\underline{R}^*$ is in general nonsymmetric.

It was remarked above that variance of a principal component, hence its contribution to the total variance, decreases as we proceed from the first to the last principal component. In fact, if the rank of $\underline{C}$ is less than q, that is, if $\underline{C}$ is positive semi-definite, at least one principal component will have zero variance. It may be desirable, for purposes of classificatory analysis, to ignore those principal components that contribute little to total variance and retain those components that have relatively large variances. The program will allow this in two ways:

(i) The user directly specifies a number, say p_1 ($\leqq q$), indicating the maximum number of principal components to be retained.

(ii) The user specifies a limit, say α, for the proportion of cumulative variance associated with principal components to be retained. In this situation, p_2 ($\leqq q$), the maximum number of components to meet this requirement, is determined in such a way that the relationship

$$\sum_{i=1}^{p_2} \lambda_i / \text{tr } \underline{C} \leqq \alpha$$

will hold. Both of the specifications (i) and (ii) must be given. The actual number of principal components to be retained, which we designate as p, is then taken as the smaller of p_1 and p_2. Clearly, p satisfies requirements set in both (i) and (ii). the last (q-p) principal components are thus eliminated for purposes of classificatory analysis.

If we now let $\underline{A}^*$ (qxp) be the matrix obtained from $\underline{A}$ by dropping the last (q-p) columns, then the transformation

$$\underline{X}^{(c)} = \frac{1}{\sqrt{n-1}}\ \underline{F}\ \underline{A}^*\ \underline{D}^{-1}(\sqrt{\lambda_i})\ ,\ i = 1, \ldots, p, \tag{9}$$

orthonormalizes the $\underline{E}$ matrix and yields a new set of $x^{(c)}$-variables ("c" for covariance option) each with mean zero and variance $1/(n-1)$. The matrix $\underline{X}^{(c)}$ (nxp) is the score matrix with respect to these variables. Column vectors of $\underline{X}^{(c)}$ form an orthonormal set - hence the term orthonormalization.

Note that the above transformation cannot be performed if any λ_i is zero. To forestall this difficulty in the program, any principal component whose variance is less than 0.001 percent of the total variance is automatically ignored for classificatory analysis, regardless of specifications given by the user.

(2) Correlation option: The approach under this option is analogous to that of the covariance option, except that $\underline{C}$ is replaced by $\underline{R}$ and $\underline{F}$ is replaced by $\underline{Z}$. With these substitutions, everything said under the covariance option applies here. Thus, under the correlation option, the principal components become

linear combinations of the standardized z-variables rather than e-variables, and the total "variance" accounted for by all principal components is tr $\underline{R}=q$. Eigenvalues λ_i ($i = 1, \ldots, q$) and matrices $\underline{A}$, $\underline{A}^*$, $\underline{G}$, $\underline{R}^*$, and the orthonormalized score matrix under the correlation option are in general different from their counterparts under the covariance option. In particular, under the correlation option, the matrix $\underline{R}^*$ is reduced to a simpler form:

$$\underline{R}^* = \underline{A}\ \underline{D}(\sqrt{\lambda_i}) \ .$$

We designate the counterpart of $\underline{X}^{(c)}$ under the correlation option as $\underline{X}^{(r)}$ (nxp) ("r" for correlation).

It may be added that the standardized score matrix $\underline{Z}$ is not defined when any of the e-variables (or f-variables) has zero variance. When this condition is encountered, the program prints a warning message, stops the execution for that job, and moves on to the next job, if any. The way to get around this problem is to use the covariance option or to exclude the useless zero-variance variables from the input data. Two "solutions" are equivalent as far as the classificatory analysis.

We now let $\underline{X}$ (nxp) stand for the matrix $\underline{X}^{(c)}$ if principal components are extracted from the covariance matrix, and as the matrix $\underline{X}^{(r)}$ if they are extracted from the correlation matrix. Then, the program assumes that $\underline{X}$ is the initial score matrix for classificatory analysis, basing the evaluation and improvement of input classification on the matrix $\underline{X}$.

Discussion

We have noted that input data can be scale-altered and/or orthonormalized, under either the covariance or correlation option, and eigenvectors associated with discriminant functions can be normalized, or left nonnormalized, depending on the discretion of the user. It is important to investigate what these options mean in context of evaluation and improvement of a classification. In particular, inasmuch as improvement of a classification is performed in the discriminant space, one wishes to know how this improvement is related to the initial space from which discriminant functions are derived. A full discussion of these aspects lies outside the scope of this contribution. To assist the user in formulating his approach, however, a brief discussion is given below. In this connection, four theorems that bear on problems raised above are stated informally in the Appendix. To avoid complications, we shall assume in our discussion below that covariance and correlation matrices in the space of input variables are positive definite (that is all respective eigenvalues are positive), and that, in the instance of orthonormalization, all principal components are retained.

First, we may inquire about effects, if any, of scale alteration and orthonormalization on our evaluation criteria. In the situation of Λ and tr $\underline{W}^{-1}\underline{B}$ criteria, the answer to this question is very simple. Since these two criteria are invariant under all nonsingular linear transformations, and scale alteration and orthonomalization are two such transformations, Λ and tr $\underline{W}^{-1}\underline{B}$ remain unaffected by scale alteration and/or orthonomalization. In practical terms, this means that, to evaluate two classifications on the basis of these two criteria, it is unnecessary and immaterial to perform scale alteration or orthonormalization on the data. Results of evaluation would not be affected by these transformations. Scale alteration, however, may be necessary to obtain a readable intermediate output.

Effects of scale alteration and orthonormalization on the tr $\underline{W}$ criterion, however, are not so straightforward. In general, this criterion, being invariant only under orthogonal transformations, would be materially affected by scale alteration and orthonormalization, so that relative rankings of two classifications on the basis of the tr $\underline{W}$ criterion would not be the same in the space of input variables as in the space of scale-altered variables or orthonormalized variables. If we let $\underline{W}^{(1)}$ (pxp) and $\underline{W}^{(2)}$ (pxp) be within scatter matrices in the space of input variables and scale-altered variables, respectively, with positive values c_i ($i = 1, \ldots, p$) as scale-alteration constants, then it becomes evident from expression (8) that

$$\text{tr}\ \underline{W}^{(2)} = \text{tr}\ [\ \underline{D}^{-1}(c_i)\ \underline{W}^{(1)}]$$

$$= \sum_{h=1}^{m} \sum_{k=1}^{n_h} \sum_{i=1}^{p} \frac{1}{c_i} (e^{(1)}_{hki} - e^{(1)}_{h.i})^2 , \quad (10)$$

where $e^{(1)}_{hki}$ and $e^{(1)}_{h.i}$ are defined the same way as x_{hki} and $x_{h.i}$ in (1). Similarly, if we let $\underline{W}^{(e)}$ (pxp) be the within scatter matrix in the e-space ($e^{(1)}$- or $e^{(2)}$-space), and $\underline{W}^{(o)}$ (pxp) the within scatter matrix in the orthonormalized space under the covariance (correlation) option, it can be shown from expression (9) (setting $\underline{A}^* = \underline{A}$) that

$$\text{tr}\ \underline{W}^{(o)} = \frac{1}{n-1}\ \text{tr}\ [\underline{A}\ \underline{D}^{-1}(\lambda_i)\ \underline{A}'\ \underline{W}^{(e)}] . \quad (11)$$

Obviously, the relation between tr $\underline{W}^{(e)}$ and tr $\underline{W}^{(o)}$ is not a simple one. It should be noted from (10), however , that, if $c_i = c$ for all i, then tr $\underline{W}$ value in the space of scale-altered variables will be a simple proportion of tr $\underline{W}$ in the space of input variables. This means that, if input scores of two classifications are scale-altered by the same constant c, then comparison of these classifications by the tr $\underline{W}$ criterion after the scale alteration is equivalent to a similar comparison made prior to the scale alteration.

It is also of interest to inquire whether scale alteration has any effect on the tr $\underline{W}$ value computed in the orthonormalized space, and how this criterion in the orthonormalized space is affected by covariance and correlation options. Answers to these questions are found in Theorems (1) and (2) (Appendix). Theorem (1) states that, relative to the orthonormalized score matrix derived directly from input data under the covariance option, the orthonormalized score matrix derived from scale-altered data under the same option represents an orthogonal transformation, that is, a linear mapping that preserves tr $\underline{W}$. Furthermore, if orthonormalization is performed under the correlation rather than the covariance option, the said orthonormalized score matrices are equal to each other. Theorem (2) states that orthonormalized score matrices generated from a given e-space (input $e^{(1)}$-space or scale-altered $e^{(2)}$-space) under covariance and correlation options also are related orthogonally, so that tr $\underline{W}$ values in orthonormalized spaces under the two options are identical. In practical terms, these relations mean that, if one were to compare two classifications on the basis of the tr $\underline{W}$ criterion in the orthonormalized space, whether input scores have been scale-altered prior to orthonormalization, or whether the covariance or correlation option is used for orthonormalization, is immaterial and would not affect the results of evaluation. Scale alteration, however, may again be necessary to obtain readable intermediate outputs. Clearly, since any transformation that preserves tr $\underline{W}$ must also preserve Λ and tr $\underline{W}^{-1}\underline{B}$, what has just been said for the tr $\underline{W}$ criterion readily applies to the Λ and tr $\underline{W}^{-1}\underline{B}$ criteria.

Another question of interest is how improvement of a classification by reduction of tr $\underline{W}$ in the discriminant space is related to the initial space from which discriminant functions are generated. The option to normalize, or leave nonnormalized, the eigenvectors associated with discriminant functions has an important bearing on this problem, as is evident from Theorems (3) and (4) (Appendix). If these vectors are left nonnormalized (Theorem (3)(ii)), it can be shown that the trace of the total scatter matrix in the discriminant space is (p + tr $\underline{W}^{-1}\underline{B}$), where p is the number of variables, so that reduction of tr $\underline{W}$ in the discriminant space is equivalent to an increase in tr $\underline{W}^{-1}\underline{B}$ (identical in initial and discriminant spaces). Thus, under this option, we may assume that improvement of a partition is tantamount to an increase in the value of tr $\underline{W}^{-1}\underline{B}$ in the initial space. This means we utilize Mahalanobis distances in the initial space to improve our partition. This is also evident from item (i) of Theorem (3).

When eigenvectors associated with discriminant functions are normalized the above relations in general do not hold, and it is difficult if not impossible to relate the improvement in the discriminant space to the initial space. This is because, under the option of normalization, tr $\underline{W}$ in the discriminant space becomes a function of lengths of nonnormalized eigenvectors of $\underline{W}^{-1}\underline{B}$, and these lengths have no simple relation to the initial space. In the special situation where the initial space is orthonormal, however, a simple solution is readily available. Theorem (4) states that, relative to an orthonormal score matrix, the discriminant score matrix under the option of normalization represents an orthogonal transformation, whereby the value of tr $\underline{W}$ is preserved. Thus, under this option, we may assume that improvement of a partition by reduction of tr $\underline{W}$ in the discriminant space is equivalent to a similar improvement by reduction of tr $\underline{W}$ in the initial (orthonormal) space. This means we utilize ordinary Euclidean distances in the initial space to improve our partition.

It is evident from foregoing remarks that improvement of a classification will be influenced by our choice to normalize, or leave nonnormalized, eigenvectors associated with discriminant functions. The improved classification in the normalized case generally will be different from its counterpart in the nonnormalized case, although differences, in most instances, will probably be small. The choice between normalization and nonnormalization is in effect equivalent to a choice between two different measures of distance. An improvement by reduction of tr $\underline{W}$ in the initial space implicitly assumes that proper measure of distance in this space is the ordinary Euclidean distance. An improvement by the tr $\underline{W}^{-1}\underline{B}$ criterion, on the other hand, assumes that the Mahalanobis distance is the proper measure of distance. These distance measures have their own merits and demerits. Since it tends to account for correlations between groups, the Mahalanobis distance has a nice intuitive appeal; it is also invariant under all nonsingular linear transformations. Its use, however, usually requires assumption of normality and equality of group covariance matrices. If these assumptions do not hold, the ordinary Euclidean metric may be a more proper measure of distance, especially if variables are uncorrelated each with an equal variance, as in the situation where input data are orthonormalized. A disadvantage of ordinary Euclidean distance is that it is invariant only under orthogonal transformations. On the other hand, its use does not require nonsingularity of the within scatter matrix.

Two final remarks, though self-evident, seem noteworthy. First, invariancy of the Mahalanobis distance (or tr $\underline{W}^{-1}\underline{B}$) under nonsingular linear transformations means that, under the option of nonnormalization, it is unnecessary and immaterial, for purposes of improvement, to perform scale alteration and orthonormalization on input data. The improved partition will not be affected by these transformations. Here again, however, scale alteration may be necessary to obtain a readable intermediate output. Second, when the option of normalization is used, ordinarily the option of orthonormalization should also be used. Otherwise reduction of tr $\underline{W}$ in the discriminant space under normalization will have no simple relation to the initial space either in terms of ordinary Euclidean distances or Mahalanobis distances.

PROGRAM DESCRIPTION

General

The ITERIM is coded in FORTRAN IV, Level H, and is to be run on an IBM System/360 computer with an available core capacity of 345 bytes or larger. It was developed and tested on the IBM 360/67 Model at the Stanford Computation Center, where compilation time was of the order of 25-30 seconds. The program makes use of the dynamic storage allocation feature of FORTRAN and avoids certain potentially troublesome features such as nonstandard returns. The entire package consists of one main driving program and four subroutines (EXEC, ORTHON, NROOT, and DATA). In addition, however, subroutines CORRE, ARRAY, and EIGEN, provided in the IBM Scientific Subroutine Package, must be available as library routines in the system. If a particular installation lacks this feature, then these subroutines, whose listings can be found in the IBM "Programmer's Manual H20-0205-3, System/360 Scientific Subroutine Package (360A-CM-03X), Version III (1968)," should be appended to the program.

Limitations on Data

For a given classification, a maximum of 300 items, 30 input variables, 25 groups, and 25 iterations are allowed. The actual number of variables used or retained for classificatory analysis cannot exceed the difference between the total number of items and number of groups. If input data fail to satisfy this requirement, then the number of variables may be reduced through orthonormalization before performing classificatory analysis. More than one classification can be processed in one run, with each classification being treated as an independent data set.

Storage Readjustment

In order to use the program on machines with less than 345 available core capacity, or to exceed data limitations noted, it is necessary to adjust the storage requirement of the program. With the dynamic storage allocation feature built into the program, this readjustment can be made in a very simple way. It is only necessary to make suitable changes in the absolute sizes of the arrays (DIMENSION statements) in the main driving program; the DIMENSION statement in the subroutines need not be, and should not be, altered. With these readjustments the program can handle almost any number of items, input variables, groups, and iterations, the avialable core capacity of the machine permitting. The way to adjust for, say, a given number of items becomes readily apparent from comparison of DIMENSION statements in the main driving program and the EXEC subroutine. In this connection, the user will find the following information on subroutine arguments useful: MAXRO = maximum number of items; KVAR = number of input variables; N = number of groups; MAXIT = maximum number of iterations requested; NSYM = (KVAR) . (KVAR + 1)/2; KK1 = (MAXRO). (KVAR); KK2 = $(\mathrm{KVAR})^2$. Input and output formats are flexible to accommodate almost any number of items, variables, groups, and iterations, so there is no need to make readjustments on them.

Computational Options

The program allows the following computational options:

1. Scale alteration of input variables through division by square roots of a set of positive constants supplied by the user. This option allows the user to suppress the scales of his input variables so that results will be printed or punched in fields specified by output formats. As a rule-of-thumb, input variables should not have variances greatly in excess of 1/n, where n is the total number of items. If input scores fail to meet this requirement, then their scales should be readjusted by the option provided. However, if the data are orthonormalized, such scale readjustment is not necessary, unless the covariance option is used for extracting principal components and absolute values of covariances are of the order of a billion or larger. It may be added, that if a scale reduction is to be done by powers of 10, this can be achieved through the input format as well.

2. Orthonormalization. This option is provided to obtain a set of uncorrelated, equal-variance variables for purposes of evaluation and improvement of a classification. When this option is used, the user must also specify
   (i) whether principal components are to be extracted from the covariance or correlation matrix,
   (ii) the maximum number of principal components to be retained for classificatory analysis, and
   (iii) the limit, in percentage, set for the cumulative variance associated with principal components to be retained for classificatory analysis.

Unless there are reasons to do otherwise, the number in (ii) should be set equal to the number of input variables, and the limit in (iii) should be set to 100 percent. The actual number of principal components retained for classificatory analysis satisfies conditions set in both (ii) and (iii).

3. The maximum number of iterations performed for the purpose of improvement of a classification. If the program is used for purposes other than improvement of a classification, this number should be set to 1.

4. Normalization of eigenvectors associated with discriminant functions. This option allows the user to have some control on the distance measure used for improvement of a classification. Ordinarily, when this option is used, the option of orthonormalization should also be used. The option of normalization applies to all iterations.

5. Computation of the F-statistic for the one-way analysis of variance. The option applies to all iterations.

6. Computation of correlations between input variables and principal components (applies if orthonormalization is requested), and between input variables and discriminant functions. The option, so far as it concerns discriminant functions, applies to all iterations.

In addition to computational options indicated above, the program provides options that concern output.

Output

Assuming that appropriate computational options are specified, a full print output from the program includes the following:

1. Results that concern orthonormalization, including the grand means, grand standard deviations, grand covariance or correlation matrix, eigenvalues and eigenvectors associated with principal components, sum of eigenvalues, and correlations between input variables and principal components.

2. The initial score matrix which serves as the basis of classificatory analysis. This is either the input score matrix, or the scale-altered score matrix, or the orthonormalized score matrix, depending on options specified by the user.

With the exception noted in (4), the print output listed in (3) through (12) below is given for each iteration.

3. The group means, group standard deviations, and grand standard deviations in the initial space.

4. The between, within, and total scatter matrices in the initial space, including appropriate degrees of freedom for each. The total scatter matrix is given only in the first iteration since it remains unchanged during iterations.

5. Traces of between and within scatter matrices in the initial space, and the ratio of these traces. These values can be used to evaluate a classification.

6. Eigenvalues and eigenvectors associated with discriminant functions. The sum of eigenvalues, also given, can be used to evaluate a classification.

7. Wilks' Lambda, and the ratio of within to total scatter in 1, ..., p-dimensional discriminant space, where p is the number of discriminant functions. Wilks' Lambda is another criterion that can be used to assess a classification.

8. The F-value and associated degrees of freedom.

9. Discriminant scores for all items, as well as the group means and grand means in the discriminant space.

10. Correlations between input variables and discriminant functions.

11. Ordinary Euclidean distances, in the discriminant space, between each item and the center of gravity of each group. Items are shown in those groups to which they were assigned during the immediately preceding iteration, thus indicating the improved classification at the end of the said iteration. The group to which a given item is assigned during the current iteration is marked as the "rank" of that item.

12. The number of core items, and its ratio to the total number of items, at the end of the current iteration.

13. A summary table that shows, for each item, groups to which it was allocated during all iterations performed. Items are arranged as in the input classification.

When the number of items and number of iterations performed are large, the print output can be voluminous. The program allows an option to suppress this output. When a full print is not requested, the total, between and within scatter matrices, and discriminant scores for items, are not printed.

In addition to print output, the ITERIM program allows the user to obtain some punch output which he may later wish to use as input to another program. The two options provided in this connection are:

1. Card output of item discriminant scores obtained at the last iteration.

2. Card output of eigenvectors and group discriminant scores obtained at the last iteration.

In either case above, each output card contains up to 8 numbers.

Basic Executional Steps

The program treats each classification as an independent data set and operates upon it, then moves on to the next data set, and repeats the operation. For each classification (data set) the basic steps of execution, given in sequence of execution, are as follows:

1. Control cards and input data are read in, and the job is annotated.

2. If option given, scale alteration is per-

formed on the input data; otherwise step (2) is skipped.

3. If option given, the input data matrix, or its rescaled form, whichever the case may be, is orthonormalized, with due cognizance to specifications given by the user. If orthonormalization is not requested, step (3) is skipped.

Now let X be the score matrix obtained by orthonormalization if this option is requested, the rescaled data matrix if scale alteration is requested but orthonormalization not requested, and the input score matrix if neither of these options is requested.

4. The matrix X is treated as the initial score matrix and the input classification identified by X is evaluated and improved, while at the same time discriminant functions and F-values based on X are computed. This process is repeated iteratively, with each iteration yielding a new, improved classification, which is evaluated and further improved at the next iteration.

5. When improvement of a classification is no longer possible, or when the maximum number of iterations specified by the user is exceeded, iterations are terminated and their history is summarized in a table.

Input Instructions

Instructions for input to the program (excluding System or "Job Control Language" cards) are described. All integers must be right-justified in their fields. Floating-point numbers must not contain more than 7 digits.

1. Multiple job card

Col. 1-5	Number of classifications (data sets) to be processed (integer).
Note:	Steps (2) through (7) below refer to a given classification (data set).

2. Data set title card

Col. 1-80	Any string of alphanumeric characters (including blanks), intended for job identification.

3. Data set control card

Col. 1-5	Total number of items (integer).
Col. 6-10	Number of input variables (integer).
Col. 11-15	Number of groups (integer).
Col. 16-20	Maximum number of iterations (integer).
Col. 25	Whether scale alteration of input variables is desired: punch 1 if yes, leave blank otherwise.
Col. 30	Whether normalization of eigenvectors associated with discriminant functions is desired: punch 1 if yes, leave blank otherwise.
Col. 35	Whether full print is desired: punch 1 if yes, leave blank otherwise.
Col. 40	Whether computation of F-statistic is desired: punch 1 if yes, leave blank otherwise.
Col. 45	Whether computation of correlations between input variables and principal components (applies if orthonormalization is requested), and between input variables and discriminant functions is desired: punch 1 if yes, leave blank otherwise.
Col. 50	Whether punch output of item discriminant scores of the last iteration is desired: punch 1 if yes, leave blank otherwise.
Col. 55	Whether punch output of eigenvectors and group discriminant scores of the last iteration is desired: punch 1 if yes, leave blank otherwise.
Col. 60	Whether orthonormalization is desired: punch 1 if yes, leave blank otherwise.
Note:	The following three options apply only when orthonormalization is requested. If this is not requested, then cols. 61-80 should be left blank.
Col. 61-65	Maximum number of principal components to be retained for classificatory analysis (integer).
Col. 70	Whether principal components are to be extracted from the covariance matrix (punch 1), or from the correlation matrix (leave blank).
Col. 71-80	Limit, in percent (e.g., 95.0), set for the cumulative variance of principal components to be retained for classificatory analysis (floating-point number).

4. Group size card(s)

Col. 1-3	Number of items in first group (integer).
Col. 4-6	Number of items in second group (integer).

. . .
. . .

Continue until all group sizes are indicated in consecutive order, allowing three columns for each group size. If DIMENSION statements are readjusted to accommodate more than 25 groups, then as many group cards as necessary, with the format given above, should be used. Note, however, that each group size card can contain up to 25 numbers (i.e., cols. 76-80 must not be used).

5. Scale alteration card(s) (optional)

To be supplied only if scale alteration of input variables is requested (1 in col. 25 of data set control card). If this transformation is not requested, then the group size card(s) is (are) immediately followed by the format card.

Col. 1-10 Scale alteration constant for the first input variable (floating-point number).

Col. 11-20 Scale alteration constant for the second input variable (floating-point number).

. . .

. . .

Continue until the scale alteration constants for all input variables are indicated in consecutive order, allowing 10 columns for each constant, up to 8 constants per card, and using as many cards as necessary.

6. Format card

Col. 1-80 Format, enclosed in parentheses, for reading in the input data matrix. The first nonblank fields must be designated as A4, A2 for item identification. Scores in the input data are to be regarded as real numbers even if, in reality, they are partly or wholly integer numbers. Hence, for example, the format (5X, A4, A2, 7I9) is not valid, while the format (5X, A4, A2, 7F9.0) is.

7. Annotated input data matrix

Each row of this matrix must contain (i) item name or item index (a string of 6 alphanumeric characters), followed by (ii) input scores for that item. The matrix is read in row-wise (item-wise). Scores of any item may be placed on more than one card, but each item must start on a new card. Items are assumed to have been arranged into groups.
Important: Each group, including the last one, must be followed by a comments-card in the data matrix. What, if anything, this card contains is immaterial.

8. If another classification (data set) is to be processed, repeat steps (2) through (7) for that classification.

Sample Problem

A full print output from a hypothetical sample problem involving 35 items, 5 groups (subsets), and 4 input variables is shown in Table 1. The listing of input to the program (excluding the System or "Job Control Language" cards) is given in Table 2. Except for item names, reshuffing of items, and their arrangement into groups, the input score matrix is the same as that given in Dixon (1967, p. 155) for principal component analysis. The input data was orthonormalized using the correlation option prior to classificatory analysis, and eigenvectors associated with discriminant functions were normalized. All principal components were retained. The rest of the options are noted in the heading of the print output in Table 1.

It will be noted that input classification, as identified by the orthonormal score matrix, was evaluated (by the tr $\underline{W}$, Wilks' Lambda, and tr $\underline{W}^{-1}\underline{B}$ criteria) and improved during the first iteration. The improved classification was then re-evaluated and further improved at the second iteration, and so on. Improved classifications were "stabilized" at the fourth iteration, whereupon iterative procedures and execution were automatically terminated. Since the initial space was orthonormal, and eigenvectors associated with discriminant functions were normalized, improvement of a partition by reduction of tr $\underline{W}$ in the discriminant space was equivalent to a similar improvement in the initial space (see "Discussion"). This is demonstrated by progressive decrease, in the initial space, of tr $\underline{W}$ from 2.08 at the first iteration to 1.99 at the fourth iteration. Interestingly, this improvement relative to the tr $\underline{W}$ criterion in the initial space was paralleled by improvements relative to Wilks' Lambda (progressively decreasing) and tr $\underline{W}^{-1}\underline{B}$ (progressively increasing) criteria as well. The number of core items (= 30) at the end of the fourth iteration indicates that "stabilization" of input classification involved displacement of 5 items from their original groups.

Listing of FORTRAN IV Program

```
C     ITERIM = "ITERATIVE IMPROVEMENTS" PROGRAM FOR EVALUATION AND
C     IMPROVEMENT OF A CLASSIFICATION. PROGRAM ACCEPTS A DATA SET
C     PARTITIONED INTO AN ARBITRARY NUMBER OF SUBSETS, EVALUATES THIS
C     PARTITION, AND IMPROVES UPON IT ITERATIVELY. IN ADDITION, A LINEAR
C     DISCRIMINANT ANALYSIS AND A ONE-WAY MULTIVARIATE ANALYSIS OF
C     VARIANCE ON THE DATA IS PERFORMED DURING EACH ITERATION.
C     IMPROVEMENT OF A PARTITION IS ACHIEVED THROUGH REDUCTION OF POOLED
C     WITHIN-SUBSETS SUM OF SQUARES BY THE NEAREST-NEIGHBOR ALGORITHM IN
C     THE DISCRIMINANT SPACE. THE FOLLOWING THREE MEASURES, WHICH MAY
C     BE USED TO ASSESS THE QUALITY OF A PARTITION, ARE COMPUTED FOR
C     A GIVEN CLASSIFICATION: (1) POOLED WITHIN-SUBSETS SUM OF SQUARES
C     (IN THE INITIAL SPACE), (2) RATIO OF WITHIN-SUBSETS SCATTER TO
C     TOTAL SCATTER (IN THE SENSE OF S.S. WILKS, 1960,1962), AND (3)
C     SUM OF EIGENVALUES ASSOCIATED WITH LINEAR DISCRIMINANT FUNCTIONS.
C     AN OPTION IS PROVIDED TO ORTHONORMALIZE INPUT DATA THROUGH
C     PRINCIPAL COMPONENTS, USING EITHER THE CORRELATION MATRIX OR THE
C     COVARIANCE MATRIX AS A BASIS OF ORTHOGONALIZATION. THE NUMBER OF
C     PRINCIPAL COMPONENTS TO BE RETAINED FOR CLASSIFICATORY ANALYSIS
C     CAN BE CONTROLLED BY SPECIFYING A MAXIMUM NUMBER AND/OR SETTING
C     A LIMIT ON THE CUMULATIVE VARIANCE. IF REQUESTED, CORRELATIONS
C     BETWEEN INPUT VARIABLES AND PRINCIPAL COMPONENTS, AND BETWEEN
C     INPUT VARIABLES AND DISCRIMINANT FUNCTIONS, THE LATTER DURING
C     EACH ITERATION, ARE COMPUTED. PROGRAM ALSO ALLOWS AN OPTION TO
C     NORMALIZE EIGENVECTORS ASSOCIATED WITH DISCRIMINANT FUNCTIONS.
C     MORE THAN ONE DATA SET (JOB) CAN BE PROCESSED IN ONE RUN.
C     RESTRICTIONS : FOR EACH DATA SET (JOB), A MAXIMUM OF 300 ITEMS,
C     30 VARIABLES, 25 SUBSETS (CLASSES), AND 25 ITERATIONS ARE ALLOWED.
C     THE ACTUAL NUMBER OF VARIABLES USED OR RETAINED FOR CLASSIFICATORY
C     ANALYSIS CANNOT EXCEED THE DIFFERENCE BETWEEN THE TOTAL NUMBER OF
C     ITEMS AND THE NUMBER OF SUBSETS.
C     LIBRARY SUBROUTINES CORRE, ARRAY, AND EIGEN, PROVIDED IN THE IBM
C     SCIENTIFIC SUBROUTINE PACKAGE, ARE REQUIRED IN THE SYSTEM.
C     PROGRAM IN FORTRAN IV(H), FOR IBM S/360, BY F. DEMIRMEN, STANFORD
C     U., 1968. PROGRAM IN PART ADAPTED FROM E. CASETTI, OFFICE OF
C     NAVAL RESEARCH, GEOGRAPHY BRANCH, TECH. REPORT NO. 12, 1964.
C
      IMPLICIT INTEGER (I-N), REAL (A-H,O-Z)
      DIMENSION TITL(20),NROW(25),NORIG(25),NAME(300,2),
     1NAMALT(300,2),ID(300),IDALT(300),X(300,30),XALT(300,30),
     2XR(300,30),XOR(300,30),DIST(300,25),IRANK(300,25),IRTEMP(300),
     3ROOT(30),XROOT(30),CUM(30),SX(30),SUMTOT(30),XBAROV(30),STALL(30),
     4STRAW(30),XMRAW(30),B(30,30),W(30,30),T(30,30),RES(30,30),
     5XAVR(25,30),SUMSET(25,30),XBAR(25,30),ST(25,30),RR(465),
     6CONX(9000),CONB(900),CONW(900),CONRES(900)
C     TO ADJUST FOR STORAGE ALLOCATION, IT IS ONLY NECESSARY TO MAKE
C     CHANGES IN THE DIMENSION STATEMENTS NOTED ABOVE.
      COMMON TITL,MAXRO,KVAR,N,MAXIT,MDIV,NORM,IFULL,IFTEST,ICOR,IPUN1,
     1IPUN2,NORTH,NUMORT,NCOV,A,NSYM,KK1,KK2,NUMBER
C     READ NUMBER OF DATA SETS TO BE PROCESSED
      READ (5,100) NUMDAT
  100 FORMAT (I5)
C     PROCESS EACH DATA SET
      DO 200 NUMBER=1,NUMDAT
C     READ TITLE AND CONTROL CARDS FOR DATA SET
      READ (5,105) TITL,MAXRO,KVAR,N,MAXIT,MDIV,NORM,IFULL,IFTEST,ICOR,
     1IPUN1,IPUN2,NORTH,NUMORT,NCOV,A
  105 FORMAT (20A4/14I5,F10.0)
      NSYM=KVAR*(KVAR+1)/2
      KK1=MAXRO*KVAR
```

```
      KK2=KVAR*KVAR
      CALL EXEC (NROW,NORIG,NAME,NAMALT,ID,IDALT,X,XALT,XR,XOR,DIST,
     1IRANK,IRTEMP,ROOT,XROOT,CUM,SX,SUMTOT,XBAROV,STALL,STRAW,XMRAW,
     2B,W,T,RES,XAVR,SUMSET,XBAR,ST,RR,CONX,CONB,CONW,CONRES)
  200 CONTINUE
      STOP
      END
C     ****************************************************************
      SUBROUTINE EXEC (NROW,NORIG,NAME,NAMALT,ID,IDALT,X,XALT,XR,XOR,
     1DIST,IRANK,IRTEMP,ROOT,XROOT,CUM,SX,SUMTOT,XBAROV,STALL,STRAW,
     2XMRAW,B,W,T,RES,XAVR,SUMSET,XBAR,ST,RR,CONX,CONB,CONW,CONRES)
C     ****************************************************************
C     SUBROUTINE EXEC TO PERFORM THE PRINCIPAL OPERATIONS.
      DIMENSION TITL(20),FMT(20),NROW(N),NORIG(N),NAME(MAXRO,2),
     1NAMALT(MAXRO,2),ID(MAXRO),IDALT(MAXRO),X(MAXRO,KVAR),
     2XALT(MAXRO,KVAR),XR(MAXRO,KVAR),XOR(MAXRO,KVAR),DIST(MAXRO,N),
     3IRANK(MAXRO,MAXIT),IRTEMP(MAXRO),ROOT(KVAR),XROOT(KVAR),CUM(KVAR),
     4SX(KVAR),SUMTOT(KVAR),XBAROV(KVAR),STALL(KVAR),STRAW(KVAR),
     5XMRAW(KVAR),B(KVAR,KVAR),W(KVAR,KVAR),T(KVAR,KVAR),RES(KVAR,KVAR),
     6XAVR(N,KVAR),SUMSET(N,KVAR),XBAR(N,KVAR),ST(N,KVAR),RR(NSYM),
     7CONX(KK1),CONB(KK2),CONW(KK2),CONRES(KK2)
      COMMON TITL,MAXRO,KVAR,N,MAXIT,MDIV,NORM,IFULL,IFTEST,ICOR,IPUN1,
     1IPUN2,NORTH,NUMORT,NCOV,A,NSYM,KK1,KK2,NUMBER
      READ (5,100) NROW
  100 FORMAT (25I3)
      IF (MDIV .GE. 1) READ (5,105) SX
  105 FORMAT (8F10.0)
      READ (5,110) FMT
  110 FORMAT (20A4)
C     ANNOTATE THIS JOB
      WRITE (6,115) NUMBER,TITL,MAXRO,KVAR,N,MAXIT,NROW
  115 FORMAT ('1ITERIM = "ITERATIVE IMPROVEMENTS" PROGRAM BY F. DEMIRMEN
     1, GEOLOGY DEPT., STANFORD U., 1968'///
     2' JOB NO.   :',I3/' JOB TITLE :',2X,20A4/
     3' NUMBER OF ITEMS           :',I5/
     4' NUMBER OF INPUT VARIABLES :',I5/
     5' NUMBER OF SUBSETS         :',I5/
     6' MAXIMUM NO. OF ITERATIONS REQUESTED :',I4/
     7' SUBSET SIZES, IN CONSECUTIVE ORDER  :',23I4/(38X,23I4))
      WRITE (6,120) FMT
  120 FORMAT (' INPUT FORMAT          :',20A4)
      IF (MDIV .GE. 1) GO TO 122
      WRITE (6,121)
  121 FORMAT (' SCALE ALTERATION OF INPUT VARIABLES NOT REQUESTED')
      GO TO 125
  122 WRITE (6,123) SX
  123 FORMAT (' SCALE ALTERATION OF INPUT VARIABLES THROUGH DIVISION BY
     1THE SQUARE ROOTS OF FOLLOWING CONSTANTS REQUESTED :'/(1X,10F13.3))
      WRITE (6,124)
  124 FORMAT (' IF ORTHONORMALIZATION IS NOT REQUESTED THE RE-SCALED DAT
     1A WILL HEREAFTER BE CALLED INITIAL DATA'/' IF ORTHONORMALIZATION I
     2S REQUESTED IT WILL BE BASED ON THE RE-SCALED DATA')
  125 IF (NORM .LE. 0) WRITE (6,126)
  126 FORMAT (' NORMALIZATION OF EIGENVECTORS ASSOCIATED WITH DISCRIMINA
     1NT FUNCTIONS NOT REQUESTED')
      IF (NORM .GE. 1) WRITE (6,127)
  127 FORMAT (' NORMALIZATION OF EIGENVECTORS ASSOCIATED WITH DISCRIMINA
     1NT FUNCTIONS REQUESTED')
      IF (IFULL .LE. 0) WRITE (6,128)
  128 FORMAT (' FULL PRINT NOT REQUESTED')
```

```
      IF (IFULL .GE. 1) WRITE (6,129)
  129 FORMAT (' FULL PRINT REQUESTED')
      IF (IFTEST .LE. 0) WRITE (6,130)
  130 FORMAT (' COMPUTATION OF F-STATISTIC NOT REQUESTED')
      IF (IFTEST .GE. 1) WRITE (6,131)
  131 FORMAT (' COMPUTATION OF F-STATISTIC REQUESTED')
      IF (ICOR .LE. 0) WRITE (6,132)
  132 FORMAT (' COMPUTATION OF CORRELATIONS NOT REQUESTED')
      IF (ICOR .GE. 1) WRITE (6,133)
  133 FORMAT (' COMPUTATION OF CORRELATIONS REQUESTED')
      IF (IPUN1 .LE. 0) WRITE (6,134)
  134 FORMAT (' PUNCH OUTPUT OF ITEM DISCRIMINANT SCORES ASSOCIATED WITH
     1 LAST ITERATION NOT REQUESTED')
      IF (IPUN1 .GE. 1) WRITE (6,135)
  135 FORMAT (' PUNCH OUTPUT OF ITEM DISCRIMINANT SCORES ASSOCIATED WITH
     1 LAST ITERATION REQUESTED')
      IF (IPUN2 .LE. 0) WRITE (6,136)
  136 FORMAT (' PUNCH OUTPUT OF SUBSET DISCRIMINANT SCORES AND EIGENVECT
     1ORS ASSOCIATED WITH LAST ITERATION NOT REQUESTED')
      IF (IPUN2 .GE. 1) WRITE (6,137)
  137 FORMAT (' PUNCH OUTPUT OF SUBSET DISCRIMINANT SCORES AND EIGENVECT
     1ORS ASSOCIATED WITH LAST ITERATION REQUESTED')
      IF (NORTH .LE. 0) WRITE (6,138)
  138 FORMAT (' ORTHONORMALIZATION NOT REQUESTED')
C     READ IN DATA
      K=0
      DO 140 M=1,N
      NR=NROW(M)
      DO 139 I=1,NR
      K=K+1
  139 READ (5,FMT) (NAME(K,J),J=1,2),(X(K,J),J=1,KVAR)
C     SKIP COMMENTS-CARD THAT FOLLOWS SUBSET
  140 READ (5,141)
  141 FORMAT (1X)
C     IF REQUESTED, DIVIDE INPUT VARIABLES BY A GIVEN SET OF CONSTANTS
      IF (MDIV .LE. 0) GO TO 151
      DO 145 I=1,KVAR
  145 SX(I)=SQRT(SX(I))
      DO 150 J=1,KVAR
      DO 150 K=1,MAXRO
  150 X(K,J)=X(K,J)/SX(J)
C     STORE RAW DATA, RAW MEANS, AND RAW ST. DEV'S IN XOR(), XMRAW(),
C     AND STRAW() IF CORRELATIONS WITH DISCR. FUNCTIONS TO BE COMPUTED.
  151 IF (ICOR .LE. 0) GO TO 155
      DO 152 K=1,MAXRO
      DO 152 J=1,KVAR
  152 XOR(K,J)=X(K,J)
      IF (NORTH .GE. 1) GO TO 155
      CALL CORRE (MAXRO,KVAR,1,XOR,XMRAW,STRAW,B,RR,CUM,ROOT,XROOT)
  155 NVAR=KVAR
      KSTOP=0
C     ORTHONORMALIZE IF OPTION GIVEN
      IF (NORTH .GE. 1) CALL ORTHON (X,XR,XMRAW,STRAW,ROOT,XROOT,CUM,
     1RES,B,RR,MAXRO,KVAR,NVAR,ICOR,NUMORT,NCOV,A,NSYM,KSTOP)
      IF (KSTOP .GE. 1) GO TO 750
C     PRINT DATA THAT WILL SERVE THE BASIS OF SUCCEEDING COMPUTATIONS
      WRITE (6,160) TITL
  160 FORMAT ('1',20A4//)
      IF (NORTH .LE. 0) WRITE (6,161)
  161 FORMAT (' INITIAL DATA'/1X,12('-')//)
```

```
      IF (NORTH .GE. 1) WRITE (6,162)
  162 FORMAT (' ORTHONORMALIZED "INITIAL" DATA'/1X,30('-')//)
      K=0
      DO 165 M=1,N
      WRITE (6,163) M
  163 FORMAT (' SUBSET',I3)
      NR=NROW(M)
      DO 165 I=1,NR
      K=K+1
  165 WRITE(6,170) K,(NAME(K,J),J=1,2),(X(K,J),J=1,NVAR)
  170 FORMAT (2X,I3,2X,A4,A2,5X,14F8.4/(18X,14F8.4))
      DO 175 K=1,MAXRO
  175 ID(K)=K
      DO 180 M=1,N
  180 NORIG(M)=NROW(M)
C     START ITERATING
      DO 660 IT=1,MAXIT
      WRITE (6,185) IT
  185 FORMAT (//' ITERATION',I3/1X,12('-'))
C     COMPUTE AND PRINT SUBSET MEANS AND GRAND MEANS
      DO 200 J=1,NVAR
      SUMTOT(J)=0.0
      K=0
      DO 195 M=1,N
      NR=NROW(M)
      SUMSET(M,J)=0.0
      DO 190 I=1,NR
      K=K+1
  190 SUMSET(M,J)=SUMSET(M,J)+X(K,J)
      XBAR(M,J)=SUMSET(M,J)/NR
  195 SUMTOT(J)=SUMTOT(J)+SUMSET(M,J)
  200 XBAROV(J)=SUMTOT(J)/MAXRO
      WRITE (6,205)
  205 FORMAT (//' MEANS OVER SUBSETS (IN INITIAL SPACE)'//)
      DO 210 M=1,N
  210 WRITE (6,215) M,(XBAR(M,J),J=1,NVAR)
  215 FORMAT (' SUBSET',I3,8X,14F8.4/(18X,14F8.4))
      WRITE (6,220) (XBAROV(J),J=1,NVAR)
  220 FORMAT (' GRAND',12X,14F8.4/(18X,14F8.4))
C     COMPUTE SUBSET STANDARD DEVIATIONS
      DO 235 J=1,NVAR
      K=0
      DO 235 M=1,N
      NR=NROW(M)
      ST(M,J)=0.0
      DO 225 I=1,NR
      K=K+1
  225 ST(M,J)=ST(M,J)+X(K,J)**2
      ST(M,J)=ST(M,J)-NR*(XBAR(M,J)**2)
      IF (ST(M,J) .GT. 0.0) GO TO 230
      ST(M,J)=0.0
      GO TO 235
  230 ST(M,J)=SQRT(ST(M,J)/(NR-1))
  235 CONTINUE
C     DEVELOP B AND W MATRICES
      DO 240 I=1,NVAR
      DO 240 J=I,NVAR
      B(I,J)=0.0
      DO 240 M=1,N
      B(I,J)=B(I,J)+NROW(M)*(XBAR(M,I)-XBAROV(I))*(XBAR(M,J)-XBAROV(J))
```

```
  240 B(J,I)=B(I,J)
      DO 245 I=1,NVAR
      DO 245 J=I,NVAR
      W(I,J)=0.0
      K=0
      DO 245 M=1,N
      NR=NROW(M)
      DO 245 L=1,NR
      K=K+1
      W(I,J)=W(I,J)+(X(K,I)-XBAR(M,I))*(X(K,J)-XBAR(M,J))
  245 W(J,I)=W(I,J)
C     COMPUTE T MATRIX AND GRAND STANDARD DEV'S ONLY IN FIRST ITERATION
      IF (IT .GT. 1) GO TO 260
      DO 250 I=1,NVAR
      DO 250 J=1,NVAR
  250 T(I,J)=B(I,J)+W(I,J)
      DO 255 I=1,NVAR
  255 STALL(I)=SQRT(T(I,I)/(MAXRO-1))
C     PRINT STANDARD DEVIATIONS
  260 WRITE (6,265)
  265 FORMAT (//' STANDARD DEVIATIONS OVER SUBSETS (IN INITIAL SPACE)'//
     1)
      DO 270 M=1,N
  270 WRITE (6,215) M,(ST(M,J),J=1,NVAR)
      WRITE (6,220) (STALL(J),J=1,NVAR)
C     PRINT B, W, AND T MATRICES IF OPTION GIVEN
C     T MATRIX TO BE PRINTED ONLY IN FIRST ITERATION
      IF (IFULL .LE. 0) GO TO 310
      M=N-1
      WRITE (6,275) M
  275 FORMAT (//' B = BETWEEN-SUBSETS SCATTER MATRIX (IN INITIAL SPACE)'
     1/' DEGREES OF FREEDOM = ',I3//)
      DO 280 I=1,NVAR
  280 WRITE (6,285) (B(I,J),J=1,NVAR)
  285 FORMAT (18X,14F8.4)
      M=MAXRO-N
      WRITE (6,290) M
  290 FORMAT (//' W = WITHIN-SUBSETS SCATTER MATRIX (IN INITIAL SPACE)'
     1/' DEGREES OF FREEDOM = ',I3//)
      DO 295 I=1,NVAR
  295 WRITE (6,285) (W(I,J),J=1,NVAR)
      IF (IT .GT. 1) GO TO 310
      M=MAXRO-1
      WRITE (6,300) M
  300 FORMAT (//' T = TOTAL SCATTER MATRIX (IN INITIAL SPACE)'
     1/' DEGREES OF FREEDOM = ',I3//)
      DO 305 I=1,NVAR
  305 WRITE (6,285) (T(I,J),J=1,NVAR)
C     COMPUTE AND PRINT TRACES OF B AND W MATRICES
  310 SUM1=0.0
      SUM2=0.0
      DO 315 I=1,NVAR
      SUM1=SUM1+B(I,I)
  315 SUM2=SUM2+W(I,I)
      SUM=SUM1/SUM2
      WRITE (6,320) SUM1,SUM2,SUM
  320 FORMAT (//' TRACE OF B = ',F10.4,' :',5X,'TRACE OF W = ',F10.4/
     1' (TRACE OF B)/(TRACE OF W) = ',F12.6)
C     COMPUTE AND PRINT EIGENVALUES AND EIGENVECTORS OF W-1 * B
      CALL ARRAY (2,NVAR,NVAR,KVAR,KVAR,CONB,B)
```

```
      CALL ARRAY (2,NVAR,NVAR,KVAR,KVAR,CONW,W)
      CALL NROOT (NVAR,CONB,CONW,ROOT,CONRES,NORM)
      CALL ARRAY (1,NVAR,NVAR,KVAR,KVAR,CONRES,RES)
      WRITE (6,325)
  325 FORMAT (//' EIGENVALUES OF  W-INVERSE * B'/)
      WRITE (6,330) (ROOT(I),I=1,NVAR)
  330 FORMAT (1X,14F9.5)
      SUM=0.0
      DO 335 I=1,NVAR
  335 SUM=SUM+ROOT(I)
      WRITE (6,340) SUM
  340 FORMAT (/' TRACE OF W-INVERSE * B = ',F12.5)
      DO 345 I=1,NVAR
  345 XROOT(I)=100*ROOT(I)/SUM
      WRITE (6,350)
  350 FORMAT (/' PERCENTAGE OF TRACE DUE TO EACH EIGENVALUE'/)
      WRITE (6,355) (XROOT(I),I=1,NVAR)
  355 FORMAT (1X,F8.2,13F9.2)
      SUM=0.0
      DO 360 I=1,NVAR
      CUM(I)=SUM+XROOT(I)
  360 SUM=CUM(I)
      WRITE (6,365)
  365 FORMAT (/' CUMULATIVE PERCENTAGE OF TRACE DUE TO EACH EIGENVALUE'/
     1)
      WRITE (6,355) (CUM(I),I=1,NVAR)
      IF (NORM .LE. 0) WRITE (6,370)
  370 FORMAT (//' EIGENVECTORS OF  W-INVERSE * B, NON-NORMALIZED, AS COL
     1UMNS'//)
      IF (NORM .GE. 1) WRITE (6,375)
  375 FORMAT (//' EIGENVECTORS OF  W-INVERSE * B, NORMALIZED, AS COLUMNS
     1'//)
      DO 380 I=1,NVAR
  380 WRITE (6,285) (RES(I,J),J=1,NVAR)
C     COMPUTE AND PRINT SCATTER RATIOS IN DISCRIMINANT SPACE
      CUM(1)=1/(1+ROOT(1))
      DO 400 I=2,NVAR
      J=I-1
      SUM=CUM(J)
  400 CUM(I)=SUM*(1/(1+ROOT(I)))
      WILKS=CUM(NVAR)
      WRITE (6,405) NVAR
  405 FORMAT (//' RATIO OF WITHIN TO TOTAL SCATTER IN 1,...,',I2,'-DIMEN
     1SIONAL DISCRIMINANT SPACE'/)
      WRITE (6,410) (CUM(I),I=1,NVAR)
  410 FORMAT (1X,14E9.2)
      WRITE (6,411) WILKS
  411 FORMAT (/' WILKS LAMBDA = ',E16.8)
C     COMPUTE F-TEST STATISTIC IF OPTION GIVEN
      IF (IFTEST .LE. 0) GO TO 425
      IF (IT .GT. 1) GO TO 415
      A=FLOAT(NVAR)
      S=SQRT(((A**2)*((N-1.0)**2)-4)/((A**2)+((N-1.0)**2)-5))
      XM=(MAXRO-1)-((A+N)/2)
      XLAMB=-((A*(N-1))-2)/4
      R=(A*(N-1))/2
      F1=2*R
      F2=(XM*S)+(2*XLAMB)
      N1=IFIX(F1+0.5)
      N2=IFIX(F2+0.5)
```

```
  415 Y=WILKS**(1/S)
      F=((1-Y)/Y)*(F2/F1)
      WRITE (6,420) F,N1,N2
  420 FORMAT (//' TEST FOR HYPOTHESIS THAT SUBSET MEAN VECTORS ARE EQUAL
     1 :'/' F = ',F10.3,' ;',5X,'DEGREES OF FREEDOM : N1 = ',I4,',  N2 =
     2 ',I8/' THE TEST ASSUMES NORMALITY AND A COMMON DISPERSION MATRIX'
     3)
C     COMPUTE ITEM DISCRIMINANT SCORES, AND PRINT IF OPTION GIVEN
  425 DO 430 K=1,MAXRO
      DO 430 J=1,NVAR
      XR(K,J)=0.0
      DO 430 L=1,NVAR
  430 XR(K,J)=XR(K,J)+X(K,L)*RES(L,J)
      IF (IFULL .LE. 0) GO TO 445
      WRITE (6,435)
  435 FORMAT (//' DISCRIMINANT SCORES FOR ITEMS'//)
      K=0
      DO 440 M=1,N
      WRITE (6,163) M
      NR=NROW(M)
      DO 440 I=1,NR
      K=K+1
  440 WRITE (6,170) K,(NAME(K,J),J=1,2),(XR(K,J),J=1,NVAR)
C     COMPUTE AND PRINT SUBSET DISCRIMINANT SCORES
  445 DO 450 M=1,N
      DO 450 J=1,NVAR
      XAVR(M,J)=0.0
      DO 450 L=1,NVAR
  450 XAVR(M,J)=XAVR(M,J)+XBAR(M,L)*RES(L,J)
      WRITE (6,455)
  455 FORMAT (//' DISCRIMINANT SCORES FOR SUBSETS'//)
      DO 460 M=1,N
  460 WRITE (6,215) M,(XAVR(M,J),J=1,NVAR)
      DO 461 I=1,NVAR
      SX(I)=0.0
      DO 461 J=1,NVAR
  461 SX(I)=SX(I)+XBAROV(J)*RES(J,I)
      WRITE (6,220) (SX(I),I=1,NVAR)
C     COMPUTE CORRELATIONS WITH INPUT VARIABLES IF OPTION GIVEN
      IF (ICOR .LE. 0) GO TO 473
C     STORE MEANS AND ST. DEVIATIONS OF DISCR. FUNCTIONS TEMPORARILY
C     IN XBAROV AND SX VECTORS.
      SUM=FLOAT(MAXRO-1)
      IF (NORM .GE. 1) GO TO 463
      DO 462 I=1,NVAR
      XBAROV(I)=SX(I)
  462 SX(I)=SQRT((1+ROOT(I))/SUM)
      GO TO 466
  463 IF (NORTH .LE. 0) GO TO 465
      DO 464 I=1,NVAR
      XBAROV(I)=0.0
  464 SX(I)=SQRT(1/SUM)
      GO TO 466
  465 CALL ARRAY (2,MAXRO,NVAR,MAXRO,KVAR,CONX,XR)
      CALL CORRE (MAXRO,NVAR,1,CONX,XBAROV,SX,CONB,RR,CUM,ROOT,XROOT)
  466 M=MAXRO-1
      DO 469 I=1,KVAR
      DO 469 J=1,NVAR
      B(I,J)=0.0
      SUM=STRAW(I)*SX(J)
```

```
      IF (SUM) 469,469,467
  467 DO 468 K=1,MAXRO
  468 B(I,J)=B(I,J)+XOR(K,I)*XR(K,J)
      B(I,J)=B(I,J)-MAXRO*XMRAW(I)*XBAROV(J)
      B(I,J)=B(I,J)/(M*SUM)
  469 CONTINUE
      WRITE (6,470)
  470 FORMAT (//' CORRELATIONS BETWEEN INPUT VARIABLES AND DISCRIMINANT
     1FUNCTIONS'/' INPUT VARIABLES IN ROWS, DISCRIMINANT FUNCTIONS IN CO
     2LUMNS'//)
      DO 471 I=1,KVAR
  471 WRITE (6,472) I, (B(I,J),J=1,NVAR)
  472 FORMAT (' VARIABLE',I4,5X,14F8.4/(18X,14F8.4))
C     COMPUTE ORDINARY EUCLIDEAN DISTANCES IN DISCRIMINANT SPACE
  473 DO 480 K=1,MAXRO
      DO 480 MM=1,N
      DIST(K,MM)=0.0
      DO 475 L=1,NVAR
  475 DIST(K,MM)=DIST(K,MM)+(XR(K,L)-XAVR(MM,L))**2
  480 DIST(K,MM)=SQRT(DIST(K,MM))
C     COMPUTE RANKS BY NOTING WHICH COL. OF GIVEN ROW HAS LEAST VALUE
C     STORE RANKS IN IRTEMP VECTOR TEMPORARILY
      DO 485 K=1,MAXRO
      DISTLO=DIST(K,1)
      IRTEMP(K)=1
      DO 485 MM=2,N
      IF (DIST(K,MM) .GE. DISTLO) GO TO 485
      DISTLO=DIST(K,MM)
      IRTEMP(K)=MM
  485 CONTINUE
C     SET FORMAT CONTROL FOR PRINTING RANKS
      IF (N .GT. 8) GO TO 490
      IFMT=1
      GO TO 505
  490 IF (N .GT. 13) GO TO 495
      IFMT=2
      GO TO 505
  495 IF (N .GT. 18) GO TO 500
      IFMT=3
      GO TO 505
  500 IFMT=4
C     PRINT DISTANCE MATRIX AND CURRENT RANKS
C     FIRST PRINT TITLES
  505 WRITE (6,510) (I,I=1,N)
  510 FORMAT (//' EUCLIDEAN DISTANCES FROM SUBSET MEANS (IN DISCRIMINANT
     1 SPACE)'//(12X,23I5))
      GO TO (515,520,525,530), IFMT
  515 WRITE (6,535)
      GO TO 555
  520 WRITE (6,540)
      GO TO 555
  525 WRITE (6,545)
      GO TO 555
  530 WRITE (6,550)
  535 FORMAT (54X,'RANK')
  540 FORMAT (79X,'RANK')
  545 FORMAT (104X,'RANK')
  550 FORMAT (128X,'RANK')
C     NOW PRINT BODY
  555 K=0
```

```
      DO 605 M=1,N
      WRITE (6,163) M
      NR=NROW(M)
      DO 605 I=1,NR
      K=K+1
      WRITE (6,560) K,(NAME(K,J),J=1,2),(DIST(K,MM),MM=1,N)
  560 FORMAT (2X,I3,1X,A4,A2,1X,23F5.2/(13X,23F5.2))
      GO TO (565,570,575,580), IFMT
  565 WRITE (6,585) IRTEMP(K)
      GO TO 605
  570 WRITE (6,590) IRTEMP(K)
      GO TO 605
  575 WRITE (6,595) IRTEMP(K)
      GO TO 605
  580 WRITE (6,600) IRTEMP(K)
  585 FORMAT ('+',54X,I2)
  590 FORMAT ('+',79X,I2)
  595 FORMAT ('+',104X,I2)
  600 FORMAT ('+',128X,I2)
  605 CONTINUE
C     STORE RANKS OF CURRENT ITERATION IN IRANK MATRIX
      DO 610 K=1,MAXRO
      L=ID(K)
  610 IRANK(L,IT)=IRTEMP(K)
C     PRINT NO. OF CORE ITEMS AND ITS RATIO
      K=0
      NCORE=0
      DO 615 M=1,N
      NR=NORIG(M)
      DO 615 I=1,NR
      K=K+1
      IF (M .EQ. IRANK(K,IT)) NCORE=NCORE+1
  615 CONTINUE
      CORRAT=NCORE/FLOAT(MAXRO)
      WRITE (6,620) NCORE,CORRAT
  620 FORMAT (/' NO. OF CORE ITEMS = ',I3/' RATIO OF NO. OF CORE ITEMS T
     1O TOTAL NO. OF ITEMS = ',F5.3)
      IF (IT .EQ. 1) GO TO 630
C     STOP ITERATING IF RESULTS OF CURRENT ITERATION SAME AS IN PREVIOUS
      DO 625 K=1,MAXRO
      IF (IRANK(K,IT) .NE. IRANK(K,IT-1)) GO TO 630
  625 CONTINUE
      ITER=IT
      GO TO 665
C     RE-ARRANGE DATA INTO NEW SUBSETS AND KEEP COUNT OF ROWS IN EACH
C     NEW SUBSET.
  630 KNU=0
      DO 645 M=1,N
      NROW(M)=0
      DO 645 K=1,MAXRO
      IF (IRTEMP(K) .NE. M) GO TO 645
      NROW(M)=NROW(M)+1
      KNU=KNU+1
      DO 635 J=1,NVAR
  635 XALT(KNU,J)=X(K,J)
      DO 640 J=1,2
  640 NAMALT(KNU,J)=NAME(K,J)
      IDALT(KNU)=ID(K)
  645 CONTINUE
C     TRANSFER RE-ARRANGED DATA INTO WORKING LOCATIONS
```

```
      DO 655 K=1,MAXRO
      DO 650 J=1,NVAR
  650 X(K,J)=XALT(K,J)
      ID(K)=IDALT(K)
      DO 655 J=1,2
  655 NAME(K,J)=NAMALT(K,J)
      ITER=IT
C     CONTINUE ITERATION
  660 CONTINUE
C     TERMINATE ITERATION AND TABULATE RESULTS
  665 WRITE (6,160) TITL
      WRITE (6,670) ITER
  670 FORMAT (' SUMMARY OF',I3,' ITERATIONS'/1X,24('-')/)
      WRITE (6,675) (J,J=1,ITER)
  675 FORMAT (18X,25I4)
C     TRANSFER CURRENT ROWS BACK INTO ORIGINAL ROWS
      DO 680 K=1,MAXRO
      L=ID(K)
      DO 680 J=1,2
  680 NAMALT(L,J)=NAME(K,J)
      K=0
      DO 690 M=1,N
      WRITE (6,685) M
  685 FORMAT (/1X,'SUBSET',I3)
      NR=NORIG(M)
      DO 690 I=1,NR
      K=K+1
  690 WRITE (6,695) K,(NAMALT(K,J),J=1,2),(IRANK(K,J),J=1,ITER)
  695 FORMAT (2X,I3,2X,A4,A2,5X,25I4/(18X,25I4))
C     IF OPTION GIVEN, PUNCH OUT ITEM DISCRIMINANT SCORES OF LAST ITER.
      IF (IPUN1 .LE. 0) GO TO 710
      DO 700 K=1,MAXRO
  700 WRITE (7,705) (NAME(K,J),J=1,2),(XR(K,J),J=1,NVAR)
  705 FORMAT (A4,A2,6X,8F8.4/(12X,8F8.4))
C     IF OPTION GIVEN, PUNCH OUT EIGENVECTORS AND SUBSET DISCRIMINANT
C     SCORES OF LAST ITERATION.
  710 IF (IPUN2 .LE. 0) GO TO 750
      DO 715 I=1,NVAR
  715 WRITE (7,720) (RES(I,J),J=1,NVAR)
  720 FORMAT (12X,8F8.4)
      DO 725 M=1,N
  725 WRITE (7,730) M,(XAVR(M,J),J=1,NVAR)
  730 FORMAT ('SUBSET',I3,3X,8F8.4/(12X,8F8.4))
  750 RETURN
      END
C     ***************************************************************
      SUBROUTINE ORTHON (X,Y,BAR,STD,B,D,T,RES,COV,R,MAXRO,KVAR,NVAR,
     1ICOR,NUMORT,NCOV,A,NSYM,KSTOP)
C     ***************************************************************
C     SUBROUTINE ORTHON FOR ORTHONORMALIZATION OF A SET OF VARIABLES.
C     THE INPUT VARIABLES ARE FIRST ORTHOGONALIZED INTO PRINCIPAL
C     COMPONENTS, WHICH ARE THEN NORMALIZED TO OBTAIN A SET OF NEW
C     UNCORRELATED VARIABLES EACH WITH VARIANCE 1/(N-1), WHERE N IS
C     THE TOTAL SAMPLE SIZE.
      IMPLICIT INTEGER (I-N), REAL (A-H,O-Z)
      DIMENSION X(MAXRO,KVAR),Y(MAXRO,KVAR),BAR(KVAR),STD(KVAR),B(KVAR),
     1D(KVAR),T(KVAR),RES(KVAR,KVAR),COV(KVAR,KVAR),R(NSYM)
      WRITE (6,10) NUMORT,A
   10 FORMAT (//' ORTHONORMALIZATION REQUESTED'/' MAXIMUM NUMBER OF PRIN
     1CIPAL COMPONENTS TO BE RETAINED :',I5/
```

```
     2' CUMULATIVE VARIANCE OF PRINCIPAL COMPONENTS TO BE RETAINED NOT T
     30 EXCEED',F7.2,'  PERCENT OF TOTAL VARIANCE'/
     4' ANY PRINCIPAL COMPONENT WHOSE VARIANCE IS LESS THAN 0.001 PERCEN
     5T OF TOTAL VARIANCE TO BE AUTOMATICALLY IGNORED')
      IF (NCOV .LE. 0) WRITE (6,15)
   15 FORMAT (' PRINCIPAL COMPONENTS TO BE EXTRACTED FROM CORRELATION MA
     1TRIX')
      IF (NCOV .GE. 1) WRITE (6,20)
   20 FORMAT (' PRINCIPAL COMPONENTS TO BE EXTRACTED FROM COVARIANCE MAT
     1RIX')
      WRITE (6,25)
   25 FORMAT (' RAW DATA REFERS TO RE-SCALED DATA IF SCALE ALTERATION IS
     1 REQUESTED, TO INPUT DATA OTHERWISE')
C     COMPUTE AND PRINT GRAND MEANS, GRAND ST. DEV'S, AND GRAND
C     CORRELATIONS OR COVARIANCES FROM RAW DATA.
      CALL CORRE (MAXRO,KVAR,1,X,BAR,STD,COV,R,B,D,T)
      WRITE (6,30) BAR
   30 FORMAT ('1GRAND MEANS COMPUTED FROM RAW DATA'//(1X,10F13.2))
      WRITE (6,35) STD
   35 FORMAT (//' GRAND STANDARD DEVIATIONS COMPUTED FROM RAW DATA'//
     1(1X,10F13.2))
      IF (NCOV .GE. 1) GO TO 55
      WRITE (6,40)
   40 FORMAT (//' C = GRAND CORRELATION MATRIX (LOWER TRIANGLE) COMPUTED
     1 FROM RAW DATA'//)
      K=0
      DO 45 I=1,KVAR
      M=K+1
      K=K+I
   45 WRITE (6,50) (R(J),J=M,K)
   50 FORMAT (1X,14F9.4)
      GO TO 85
   55 M=MAXRO-1
      DO 60 I=1,KVAR
      DO 60 J=1,KVAR
   60 COV(I,J)=COV(I,J)/M
      WRITE (6,65)
   65 FORMAT (//' C = GRAND COVARIANCE MATRIX (LOWER TRIANGLE) COMPUTED
     1FROM RAW DATA'//)
      DO 70 I=1,KVAR
   70 WRITE (6,75) (COV(I,J),J=1,I)
   75 FORMAT (1X,10F13.2)
C     STORE COVARIANCE MATRIX IN IBM STORAGE MODE 1
      K=0
      DO 80 J=1,KVAR
      DO 80 I=1,J
      K=K+1
   80 R(K)=COV(I,J)
C     COMPUTE AND PRINT EIGENVALUES AND NORMALIZED EIGENVECTORS OF
C     CORRELATION MATRIX OR COVARIANCE MATRIX, WHICHEVER IS REQUESTED.
   85 CALL EIGEN (R,RES,KVAR,0)
C     RECOVER EIGENVALUES FROM THE MAIN DIAGONAL OF R MATRIX
      K=0
      DO 90 I=1,KVAR
      K=K+I
   90 D(I)=R(K)
      WRITE (6,95) D
   95 FORMAT (//' EIGENVALUES OF C MATRIX'//(1X,10F13.2))
      SUM=0.0
      DO 96 I=1,KVAR
```

```
   96 SUM=SUM+D(I)
      WRITE (6,97) SUM
   97 FORMAT (//' TRACE OF C MATRIX = ',F20.6)
      DO 98 I=1,KVAR
   98 T(I)=100*D(I)/SUM
      WRITE (6,99) T
   99 FORMAT (//' PERCENTAGE OF TRACE DUE TO EACH EIGENVALUE'//
     1(1X,14F9.2))
C     DETERMINE NO. OF PRINCIPAL COMPONENTS TO BE RETAINED
      MM=0
      SUM=0.0
      DO 100 I=1,KVAR
      B(I)=SUM+T(I)
      IF ((A .GE. B(I) .OR. B(I)-A .LT. .01) .AND. T(I) .GE. .001)
     1MM=MM+1
  100 SUM=B(I)
      IF (MM .GT. NUMORT) MM=NUMORT
      NVAR=MM
C     NVAR IS THE NUMBER OF PRINCIPAL COMPONENTS RETAINED
      WRITE (6,105) B
  105 FORMAT (//' CUMULATIVE PERCENTAGE OF TRACE DUE TO EACH EIGENVALUE'
     1//(1X,14F9.2))
      WRITE (6,110)
  110 FORMAT (//' EIGENVECTORS OF C MATRIX, NORMALIZED, AS COLUMNS'//)
      DO 115 I=1,KVAR
  115 WRITE (6,50) (RES(I,J),J=1,KVAR)
C     COMPUTE AND PRINT CORRELATIONS WITH PRIN. COMPS. IF OPTION GIVEN
      IF (ICOR. LE. 0) GO TO 121
      DO 117 I=1,KVAR
      DO 117 J=1,KVAR
      IF (STD(I) .GT. 0.0 .AND. D(J) .GT. 0.0) GO TO 116
      COV(I,J)=0.0
      GO TO 117
  116 COV(I,J)=RES(I,J)*SQRT(D(J))
      IF (NCOV .GE. 1) COV(I,J)=COV(I,J)/STD(I)
  117 CONTINUE
      WRITE (6,118)
  118 FORMAT (//' CORRELATIONS BETWEEN INPUT VARIABLES AND PRINCIPAL COM
     1PONENTS'/' INPUT VARIABLES IN ROWS, COMPONENTS IN COLUMNS'//)
      DO 119 I=1,KVAR
  119 WRITE (6,120) I, (COV(I,J),J=1,KVAR)
  120 FORMAT (' VARIABLE',I4,5X,14F8.4/(18X,14F8.4))
C     TRANSFORM INPUT VARIABLES TO HAVE ZERO MEANS
  121 DO 122 J=1,KVAR
      DO 122 K=1,MAXRO
  122 X(K,J)=X(K,J)-BAR(J)
C     IF ORTHOGONALIZATION IS BASED ON CORRELATION MATRIX, STANDARDIZE
C     INPUT VARIABLES TO HAVE UNIT VARIANCES.
      IF (NCOV .GE. 1) GO TO 126
      DO 123 J=1,KVAR
      IF (STD(J) .LE. 0) GO TO 124
      DO 123 K=1,MAXRO
  123 X(K,J)=X(K,J)/STD(J)
      GO TO 126
  124 WRITE (6,125)
  125 FORMAT (//' JOB TERMINATED : AT LEAST ONE INPUT VARIABLE HAS ZERO
     1VARIANCE AND STANDARDIZATION CANNOT BE PERFORMED'/
     2' DELETE SUCH USELESS VARIABLES OR USE COVARIANCE OPTION')
      KSTOP=1
      RETURN
```

```
C     COMPUTE DATA SCORES WITH RESPECT TO PRINCIPAL COMPONENTS, AND
C     NORMALIZE. USE AT MOST AS MANY PRINCIPAL COMPONENTS AS REQUESTED.
  126 DO 130 K=1,MAXRO
      DO 130 J=1,NVAR
      Y(K,J)=0.0
      DO 130 L=1,KVAR
  130 Y(K,J)=Y(K,J)+X(K,L)*RES(L,J)
      XM=SQRT(FLOAT(MAXRO-1))
      DO 135 J=1,NVAR
      DO 135 K=1,MAXRO
  135 X(K,J)=Y(K,J)/(XM*SQRT(D(J)))
      WRITE (6,140) NVAR
  140 FORMAT (//' NOTE : CLASSIFICATORY ANALYSIS AND ITERATIONS ARE STAR
     1TING'/' NUMBER OF PRINCIPAL COMPONENTS RETAINED :',I5/
     2' ALL COMPUTATIONS THAT FOLLOW ARE BASED ON ORTHONORMALIZED DATA'/
     3' SPACE DEFINED BY ORTHONORMAL VARIABLES WILL HEREAFTER BE CALLED
     4INITIAL SPACE')
      RETURN
      END
C     ***********************************
      SUBROUTINE NROOT (M,A,B,XL,X,NORM)
C     ***********************************
C     EIGENVALUES AND EIGENVECTORS OF A REAL MATRIX OF THE FORM
C     B-INVERSE TIMES A, WHERE A AND B ARE SYMMETRIC, B POS. DEFINITE.
C     SUBROUTINE SLIGHTLY MODIFIED FROM SUBROUTINE OF SAME NAME PROVIDED
C     IN THE IBM SCIENTIFIC SUBROUTINE PACKAGE.
      IMPLICIT INTEGER (I-N), REAL (A-H,O-Z)
      DIMENSION A(1),B(1),XL(1),X(1)
C     STORE MATRIX B IN IBM STORAGE MODE 1
      K=1
      DO 100 J=2,M
      L=M*(J-1)
      DO 100 I=1,J
      L=L+1
      K=K+1
  100 B(K)=B(L)
C     COMPUTE EIGENVALUES AND EIGENVECTORS OF B
      MV=0
      CALL EIGEN (B,X,M,MV)
C     FORM RECIPROCALS OF SQUARE ROOTS OF EIGENVALUES. THE RESULTS ARE
C     PREMULTIPLIED BY THE ASSOCIATED EIGENVECTORS.
      L=0
      DO 110 J=1,M
      L=L+J
  110 XL(J)=1.0/SQRT(ABS(B(L)))
      K=0
      DO 115 J=1,M
      DO 115 I=1,M
      K=K+1
  115 B(K)=X(K)*XL(J)
C     FORM (B**(-1/2))TRANSPOSE * A * (B**(-1/2))
      DO 120 I=1,M
      N2=0
      DO 120 J=1,M
      N1=M*(I-1)
      L=M*(J-1)+I
      X(L)=0.0
      DO 120 K=1,M
      N1=N1+1
      N2=N2+1
```

```
  120 X(L)=X(L)+B(N1)*A(N2)
      L=0
      DO 130 J=1,M
      DO 130 I=1,J
      N1=I-M
      N2=M*(J-1)
      L=L+1
      A(L)=0.0
      DO 130 K=1,M
      N1=N1+M
      N2=N2+1
  130 A(L)=A(L)+X(N1)*B(N2)
C     COMPUTE EIGENVALUES AND EIGENVECTORS OF A
      CALL EIGEN (A,X,M,MV)
      L=0
      DO 140 I=1,M
      L=L+I
  140 XL(I)=A(L)
C     COMPUTE EIGENVECTORS OF B-INVERSE * A
      DO 150 I=1,M
      N2=0
      DO 150 J=1,M
      N1=I-M
      L=M*(J-1)+I
      A(L)=0.0
      DO 150 K=1,M
      N1=N1+M
      N2=N2+1
  150 A(L)=A(L)+B(N1)*X(N2)
C     NORMALIZE EIGENVECTORS IF SO REQUESTED
      IF (NORM .LE. 0) GO TO 185
      L=0
      K=0
      DO 180 J=1,M
      SUMV=0.0
      DO 170 I=1,M
      L=L+1
      IF (ABS(A(L)) .LT. 1.0E-35) A(L)=0.0
  170 SUMV=SUMV+A(L)*A(L)
  175 SUMV=SQRT(SUMV)
      DO 180 I=1,M
      K=K+1
  180 X(K)=A(K)/SUMV
      GO TO 195
  185 K=0
      DO 190 J=1,M
      DO 190 I=1,M
      K=K+1
  190 X(K)=A(K)
  195 RETURN
      END
C     ***************
      SUBROUTINE DATA
C     ***************
C     THIS IS A DUMMY SUBROUTINE REQUIRED IN THE SUBROUTINE CORRE.
      RETURN
      END
```

Table 1.—Complete print output from hypothetical problem.

```
ITERIM = "ITERATIVE IMPROVEMENTS" PROGRAM BY F. DEMIRMEN, GEOLOGY DEPT., STANFORD U., 1968


JOB NC.    :  1
JOB TITLE :  SAMFLE PROBLEM USING HYPCTHETICAL CATA
NUMBER OF ITEMS            :   35
NUMBER OF INPUT VARIABLES :    4
NUMBER OF SUBSETS          :    5
MAXIMUM NO. CF ITERATIONS REQUESTEC :   5
SUBSET SIZES, IN CCNSECUTIVE ORCER  :   6   8   5   9   7
INPUT FORMAT          :(A4,A2,6X,4F6.0)
SCALE ALTERATION OF INPUT VARIABLES NOT REQUESTED
NORMALIZATION CF EIGENVECTORS ASSOCIATED WITH DISCRIMINANT FUNCTIONS REQUESTED
FULL PRINT REQUESTEC
COMPUTATION CF F-STATISTIC REQUESTED
COMPUTATICN CF CCRRELATIONS REQUESTED
PUNCH OUTPUT OF ITEM DISCRIMINANT SCORES ASSOCIATED WITH LAST ITERATION NOT REQUESTED
PUNCH OUTPUT OF SUBSET DISCRIMINANT SCORES AND EIGENVECTORS ASSOCIATED WITH LAST ITERATION NOT REQUESTED


ORTHONORMALIZATION REQUESTED
MAXIMUM NUMBER CF PRINCIPAL COMPONENTS TO BE RETAINED :    4
CUMULATIVE VARIANCE OF PRINCIPAL CCMPONENTS TO BE RETAINED NOT TO EXCEED 100.00  PERCENT OF TOTAL VARIANCE
ANY PRINCIPAL COMPONENT WHOSE VARIANCE IS LESS THAN 0.001 PERCENT OF TOTAL VARIANCE TO BE AUTOMATICALLY IGNORED
PRINCIPAL COMPCNENTS TO BE EXTRACTEC FRCM CORRELATION MATRIX
RAW DATA REFERS TO RE-SCALED DATA IF SCALE ALTERATICN IS REQUESTEC, TO INPUT DATA OTHERWISE
```

```
GRAND MEANS COMPUTED FROM RAW DATA

      112.29        103.09        101.69        99.60

GRANC STANCARC CEVIATIONS COMPUTED FROM RAW DATA

       51.41         21.61         26.C1         75.45

C = GRAND CORRELATICN MATRIX (LOWER TRIANGLE) COMPUTED FROM RAW DATA

   1.C000
   0.7226   1.C00C
  -0.5798  -C.6568   1.0000
   0.8000   C.8950  -0.7525   1.0000

EIGENVALUES CF C MATRIX

       3.21          0.43          0.28          0.08

TRACE CF C MATRIX =             3.999989

PERCENTAGE OF TRACE DUE TO EACH EIGENVALUE

    8C.34    1C.75    6.88     2.C3

CUMULATIVE PERCENTAGE OF TRACE CUE TO EACH EIGENVALUE

    80.34    91.09    97.97   100.00

EIGENVECTORS OF C MATRIX, NORMALIZED, AS COLUMNS

   0.4831   C.5453  -0.6577  -0.1915
   0.5127   C.1667   0.6652  -0.5166
  -0.4619   C.8196   0.2886   0.1777
   0.5389   C.0551   0.2041   0.8154

CORRELATIONS BETWEEN INPUT VARIABLES AND PRINCIPAL COMPCNENTS
INPUT VARIABLES IN ROWS, COMPONENTS IN COLUMNS

VARIABLE   1       0.8660  0.3576 -0.3451 -C.0545
VARIABLE   2       0.9190  0.1093  0.3490 -0.1471
VARIABLE   3      -C.8280  0.5375  0.1514  0.0506
VARIABLE   4       0.9661  0.0361  0.1071  C.2322

NOTE : CLASSIFICATORY ANALYSIS AND ITERATIONS ARE STARTING
NUMBER OF PRINCIPAL COMPONENTS RETAINED :     4
ALL CCMPUTATIONS THAT FOLLCW ARE BASED ON CRTHONORMALIZED DATA
SPACE CEFINEC BY ORTHONORMAL VARIABLES WILL HEREAFTER BE CALLED INITIAL SPACE
```

SAMPLE PROBLEM USING HYPOTHETICAL DATA

ORTHONORMALIZED "INITIAL" DATA

SUBSET 1					
1	S- 2	-0.1384	0.2689	0.0371	0.0214
2	S- 3	-0.0769	0.2439	-0.0009	-0.0269
3	S- 6	0.0118	0.2370	-0.0546	-0.0371
4	S- 8	-0.0377	0.1214	0.0586	-0.2178
5	S- 9	-0.0682	0.0563	0.1150	-0.3225
6	S-10	-0.0460	0.1186	0.0430	-0.1465
SUBSET 2					
7	S- 4	-0.0076	0.3022	-0.1338	-0.0280
8	S- 5	0.0300	0.3263	-0.2051	-0.1063
9	S- 7	0.0074	0.1732	0.0439	0.0141
10	S-11	0.0477	0.1104	-0.0184	-0.0367
11	S-12	0.0645	0.1215	-0.1374	-0.0438
12	S-13	0.1535	-0.0357	-0.2035	0.0226
13	S-14	0.0864	-0.0828	-0.1530	-0.2037
14	S-21	0.0700	0.0320	-0.2590	-0.2024
SUBSET 3					
15	S- 1	-0.2455	0.2677	0.0795	0.3693
16	S-17	-0.1790	-0.1671	-0.0196	0.1181
17	S-18	-0.2018	-0.2457	-0.0978	0.1634
18	S-19	-0.1833	-0.2656	-0.3098	0.1385
19	S-20	-0.0880	-0.1525	-0.3706	0.0264
SUBSET 4					
20	S-22	0.2340	-0.3161	-0.2717	-0.1685
21	S-23	0.3412	-0.0434	-0.1262	0.0784
22	S-24	0.4703	0.1274	0.0076	0.6277
23	S-25	0.3202	-0.0544	0.2228	-0.1501
24	S-26	0.1788	-0.1488	0.2967	-0.1194
25	S-27	0.1509	-0.1516	0.2769	0.0353
26	S-28	0.0966	-0.1207	0.2573	0.0501
27	S-29	0.0376	-0.0979	0.1125	0.0557
28	S-30	0.0200	0.0160	0.0909	-0.0746
SUBSET 5					
29	S-15	-0.0595	-0.1613	0.0102	-0.0315
30	S-16	-0.1474	-0.1433	0.0800	0.0839
31	S-31	-0.0521	-0.0644	0.2411	-0.0665
32	S-32	-0.1447	-0.0441	0.2184	-0.0362
33	S-33	-0.2107	-0.0851	0.0484	0.0885
34	S-34	-0.2347	-0.1015	0.0362	0.1671
35	S-35	-0.1992	-0.0410	0.0856	-0.0422

ITERATION 1

MEANS OVER SUBSETS (IN INITIAL SPACE)

SUBSET 1	-0.0592	0.1743	0.0330	-0.1215
SUBSET 2	0.0565	0.1184	-0.1333	-0.0730
SUBSET 3	-0.1795	-0.1126	-0.1437	0.1632
SUBSET 4	0.2055	-0.0877	0.0963	0.0372
SUBSET 5	-0.1498	-0.0915	0.1028	0.0233

```
GRAND                 -0.0000  0.0000  0.0000  0.0000


STANDARD DEVIATIONS OVER SUBSETS (IN INITIAL SPACE)


SUBSET  1              0.0497  0.0867  0.0571  0.1317
SUBSET  2              0.0505  0.1472  0.1006  0.0893
SUBSET  3              0.0575  0.2181  0.1913  0.1263
SUBSET  4              0.1503  0.1231  0.1961  0.2412
SUBSET  5              0.0720  0.0469  0.0906  0.0890
GRAND                  0.1715  0.1715  0.1715  0.1715


B = BETWEEN-SUBSETS SCATTER MATRIX (IN INITIAL SPACE)
DEGREES OF FREEDOM =    4


                       0.7448  0.0264  0.1273 -0.0919
                       0.0264  0.4859 -0.1527 -0.3325
                       0.1273 -0.1527  0.4093 -0.0145
                      -0.0919 -0.3325 -0.0145  0.2807


W = WITHIN-SUBSETS SCATTER MATRIX (IN INITIAL SPACE)
DEGREES OF FREEDOM =   30


                       0.2552 -0.0264 -0.1273  0.0919
                      -0.0264  0.5141  0.1527  0.3325
                      -0.1273  0.1527  0.5907  0.0144
                       0.0919  0.3325  0.0144  0.7194


T = TOTAL SCATTER MATRIX (IN INITIAL SPACE)
DEGREES OF FREEDOM =   34


                       1.0000  0.0000 -0.0000  0.0000
                       0.0000  1.0000 -0.0000  0.0000
                      -0.0000 -0.0000  1.0000 -0.0000
                       0.0000  0.0000 -0.0000  1.0000


TRACE OF B =      1.9207 ;      TRACE OF W =      2.0794
(TRACE OF B)/(TRACE OF W) =      0.923685


EIGENVALUES OF  W-INVERSE * B

  4.54475  3.06972  0.51366  0.00744

TRACE OF W-INVERSE * B =       8.13557

PERCENTAGE OF TRACE DUE TO EACH EIGENVALUE

   55.86     37.73     6.31     0.09

CUMULATIVE PERCENTAGE OF TRACE DUE TO EACH EIGENVALUE

   55.86     93.59    99.91   100.00
```

EIGENVECTORS OF W-INVERSE * B, NORMALIZED, AS COLUMNS

0.8236	-0.4391	-0.3578	0.0302
0.3956	0.6807	0.1261	0.6035
0.1220	-0.4474	0.8509	0.2469
-0.3878	-0.3790	-0.3635	0.7576

RATIO OF WITHIN TO TOTAL SCATTER IN 1,..., 4-DIMENSIONAL DISCRIMINANT SPACE

0.18E 00 0.44E-01 0.29E-01 0.29E-01

WILKS LAMBDA = 0.29060736E-01

TEST FOR HYPOTHESIS THAT SUBSET MEAN VECTORS ARE EQUAL :
F = 11.347 ; DEGREES OF FREEDOM : N1 = 16, N2 = 83
THE TEST ASSUMES NORMALITY AND A COMMON DISPERSION MATRIX

DISCRIMINANT SCORES FOR ITEMS

SUBSET 1					
1	S- 2	-0.0114	0.2190	0.1072	0.1835
2	S- 3	0.0435	0.2104	0.0673	0.1242
3	S- 6	0.1112	0.1947	-0.0072	0.1018
4	S- 8	0.1085	0.1555	0.1578	-0.0784
5	S- 9	0.1052	0.1390	0.2465	-0.1840
6	S-10	0.0711	0.1372	0.1213	-0.0301
SUBSET 2					
7	S- 4	0.1078	0.2795	-0.0628	0.1280
8	S- 5	0.1700	0.3410	-0.1054	0.0667
9	S- 7	0.0745	0.0897	0.0515	0.1263
10	S-11	0.0949	0.0764	-0.0054	0.0357
11	S-12	0.1015	0.1325	-0.1087	0.0082
12	S-13	0.0787	-0.0092	-0.2408	-0.0500
13	S-14	0.0987	0.0513	-0.0975	-0.2395
14	S-21	0.1172	0.1836	-0.1678	-0.1958
SUBSET 3					
15	S- 1	-0.2297	0.1145	0.0550	0.4536
16	S-17	-0.2617	-0.0711	-0.0167	-0.0216
17	S-18	-0.3387	-0.0968	-0.1014	-0.0548
18	S-19	-0.3476	-0.0142	-0.2819	-0.1374
19	S-20	-0.1883	0.0907	-0.3127	-0.1661
SUBSET 4					
20	S-22	0.0998	-0.1324	-0.2935	-0.3784
21	S-23	0.2181	-0.1526	-0.2634	0.0123
22	S-24	0.1952	-0.3611	-0.3740	0.5685
23	S-25	0.3276	-0.2204	0.1227	-0.0819
24	S-26	0.1708	-0.2673	0.2131	-0.1016
25	S-27	0.0844	-0.3067	0.1496	0.0082
26	S-28	0.0438	-0.2587	0.1509	0.0316
27	S-29	-0.0157	-0.1545	0.0496	0.0121
28	S-30	0.0628	-0.0103	0.0994	-0.0238
SUBSET 5					
29	S-15	-0.0994	-0.0763	0.0210	-0.1205
30	S-16	-0.2009	-0.1004	0.0722	-0.0076

31	S-31	-0.0132	-0.1037	0.2398	-0.0313
32	S-32	-0.0959	-0.0505	0.2452	-0.0045
33	S-33	-0.2356	-0.0206	0.0737	0.0212
34	S-34	-0.2938	-0.0455	0.0412	0.0672
35	S-35	-0.1535	0.0372	0.1543	-0.0416

DISCRIMINANT SCORES FOR SUBSETS

SUBSET 1	0.0714	0.1760	0.1155	0.0195
SUBSET 2	0.1054	0.1431	-0.0921	-0.0151
SUBSET 3	-0.2732	0.0046	-0.1315	0.0147
SUBSET 4	0.1319	-0.2071	-0.0162	0.0052
SUBSET 5	-0.1560	-0.0514	0.1211	-0.0167
GRAND	0.0000	0.0000	0.0000	0.0000

CORRELATIONS BETWEEN INPUT VARIABLES AND DISCRIMINANT FUNCTIONS
INPUT VARIABLES IN ROWS, DISCRIMINANT FUNCTIONS IN COLUMNS

VARIABLE 1	0.8338	0.0382	-0.5386	0.1155
VARIABLE 2	0.8997	-0.4296	0.0354	0.0685
VARIABLE 3	-0.4704	0.6426	0.4745	0.3751
VARIABLE 4	0.7329	-0.5356	-0.3344	0.2533

EUCLIDEAN DISTANCES FROM SUBSET MEANS (IN DISCRIMINANT SPACE)

	1	2	3	4	5	RANK
SUBSET 1						
1 S- 2	0.19	0.31	0.45	0.50	0.37	1
2 S- 3	0.12	0.23	0.44	0.45	0.36	1
3 S- 6	0.15	0.15	0.45	0.41	0.40	2
4 S- 8	0.11	0.26	0.51	0.41	0.34	1
5 S- 9	0.25	0.38	0.59	0.47	0.38	1
6 S-10	0.06	0.22	0.45	0.38	0.30	1
SUBSET 2						
7 S- 4	0.24	0.20	0.49	0.50	0.48	2
8 S- 5	0.30	0.22	0.56	0.56	0.56	2
9 S- 7	0.15	0.21	0.42	0.33	0.31	1
10 S-11	0.16	0.12	0.40	0.29	0.31	2
11 S-12	0.23	0.03	0.40	0.35	0.39	2
12 S-13	0.41	0.22	0.37	0.31	0.43	2
13 S-14	0.36	0.24	0.45	0.37	0.42	2
14 S-21	0.36	0.20	0.48	0.47	0.50	2
SUBSET 3						
15 S- 1	0.54	0.60	0.49	0.66	0.51	3
16 S-17	0.44	0.43	0.14	0.42	0.17	3
17 S-18	0.54	0.51	0.14	0.49	0.29	3
18 S-19	0.63	0.53	0.23	0.60	0.46	3
19 S-20	0.54	0.40	0.28	0.56	0.48	3
SUBSET 4						
20 S-22	0.65	0.50	0.58	0.48	0.61	4
21 S-23	0.52	0.36	0.53	0.27	0.55	4
22 S-24	0.92	0.83	0.85	0.69	0.90	4
23 S-25	0.48	0.48	0.70	0.26	0.52	4
24 S-26	0.48	0.52	0.64	0.26	0.41	4
25 S-27	0.48	0.51	0.55	0.20	0.35	4

```
  26 S-28       0.44 0.48 0.50 0.20 0.29                                 4
  27 S-29       0.35 0.35 0.35 0.17 0.19                                 4
  28 S-30       0.19 0.25 0.41 0.24 0.22                                 1
SUBSET  5
  29 S-15       0.35 0.34 0.28 0.30 0.16                                 5
  30 S-16       0.39 0.42 0.24 0.36 0.08                                 5
  31 S-31       0.32 0.43 0.47 0.31 0.19                                 5
  32 S-32       0.31 0.44 0.42 0.38 0.14                                 5
  33 S-33       0.37 0.41 0.21 0.42 0.10                                 5
  34 S-34       0.44 0.47 0.19 0.46 0.18                                 5
  35 S-35       0.27 0.37 0.32 0.42 0.10                                 5

NO. OF CORE ITEMS =   32
RATIO OF NO. OF CORE ITEMS TO TOTAL NO. OF ITEMS = 0.914


ITERATION  2
------------


MEANS OVER SUBSETS (IN INITIAL SPACE)


SUBSET  1             -0.0485  0.1426  0.0554 -0.1075
SUBSET  2              0.0570  0.1264 -0.1456 -0.0794
SUBSET  3             -0.1795 -0.1126 -0.1437  0.1632
SUBSET  4              0.2287 -0.1007  0.0970  0.0512
SUBSET  5             -0.1498 -0.0915  0.1028  0.0233
GRAND                 -0.0000  0.0000  0.0000  0.0000


STANDARD DEVIATIONS OVER SUBSETS (IN INITIAL SPACE)


SUBSET  1              0.0535  0.0928  0.0379  0.1284
SUBSET  2              0.0499  0.1523  0.0797  0.0839
SUBSET  3              0.0575  0.2181  0.1913  0.1263
SUBSET  4              0.1424  0.1249  0.2097  0.2539
SUBSET  5              0.0720  0.0469  0.0906  0.0890
GRAND                  0.1715  0.1715  0.1715  0.1715


B = BETWEEN-SUBSETS SCATTER MATRIX (IN INITIAL SPACE)
DEGREES OF FREEDOM =    4


                       0.7791  0.0221  0.1134 -0.0770
                       0.0221  0.4733 -0.1550 -0.3357
                       0.1134 -0.1550  0.4435 -0.0099
                      -0.0770 -0.3357 -0.0099  0.2893


W = WITHIN-SUBSETS SCATTER MATRIX (IN INITIAL SPACE)
DEGREES OF FREEDOM =   30


                       0.2209 -0.0221 -0.1134  0.0770
                      -0.0221  0.5267  0.1550  0.3357
                      -0.1134  0.1550  0.5565  0.0099
                       0.0770  0.3357  0.0099  0.7108
```

```
TRACE OF B =       1.9851 ;      TRACE OF W =       2.0149
(TRACE OF B)/(TRACE OF W) =      0.985243


EIGENVALUES OF  W-INVERSE * B

  4.92224  3.30564  0.60868  0.00794

TRACE OF W-INVERSE * B =       8.84450

PERCENTAGE OF TRACE DUE TO EACH EIGENVALUE

   55.65    37.38     6.88      0.09

CUMULATIVE PERCENTAGE OF TRACE DUE TO EACH EIGENVALUE

   55.65    93.03    99.91   100.00


EIGENVECTORS OF  W-INVERSE * B, NORMALIZED, AS COLUMNS


                    0.9078 -0.2672 -0.3225  0.0234
                    0.2517  0.7302  0.1481  0.6177
                    0.1722 -0.4283  0.8565  0.2308
                   -0.2880 -0.4604 -0.3748  0.7514


RATIO OF WITHIN TO TOTAL SCATTER IN 1,..., 4-DIMENSIONAL DISCRIMINANT SPACE

 0.17E 00 0.39E-01 0.24E-01 0.24E-01

WILKS LAMBDA =   0.24186451E-01


TEST FOR HYPOTHESIS THAT SUBSET MEAN VECTORS ARE EQUAL :
F =      12.372 ;      DEGREES OF FREEDOM : N1 =    16,  N2 =        83
THE TEST ASSUMES NORMALITY AND A COMMON DISPERSION MATRIX


DISCRIMINANT SCORES FOR ITEMS


SUBSET  1
   1  S- 2          -0.0577  0.2075  0.1082  0.1875
   2  S- 3          -0.0008  0.2114  0.0702  0.1284
   3  S- 8           0.0691  0.1739  0.1619 -0.0760
   4  S- 9           0.0650  0.1585  0.2496 -0.1826
   5  S-10           0.0377  0.1479  0.1242 -0.0279
   6  S- 7           0.0538  0.0992  0.0556  0.1279
   7  S-30           0.0593  0.0018  0.1018 -0.0247
SUBSET  2
   8  S- 6           0.0716  0.2104 -0.0015  0.1062
   9  S- 4           0.0541  0.2929 -0.0569  0.1346
  10  S- 5           0.1046  0.3670 -0.0972  0.0750
  11  S-11           0.0785  0.0927 -0.0010  0.0375
  12  S-12           0.0781  0.1506 -0.1041  0.0119
  13  S-13           0.0888  0.0097 -0.2376 -0.0484
  14  S-14           0.0899  0.0757 -0.0948 -0.2375
  15  S-21           0.0853  0.2088 -0.1638 -0.1904
```

SUBSET 3					
16	S- 1	-0.2481	0.0570	0.0485	0.4555
17	S-17	-0.2420	-0.1201	-0.0281	-0.0232
18	S-18	-0.3089	-0.1588	-0.1163	-0.0563
19	S-19	-0.3265	-0.0760	-0.2975	-0.1358
20	S-20	-0.1897	0.0588	-0.3216	-0.1619
SUBSET 4					
21	S-22	0.1346	-0.0994	-0.2918	-0.3791
22	S-23	0.2545	-0.1049	-0.2539	0.0110
23	S-24	0.2795	-0.3249	-0.3616	0.5632
24	S-25	0.3586	-0.1516	0.1358	-0.0875
25	S-26	0.2103	-0.2285	0.2192	-0.1090
26	S-27	0.1364	-0.2859	0.1528	0.0003
27	S-28	0.0872	-0.2472	0.1525	0.0247
28	S-29	0.0128	-0.1553	0.0488	0.0083
SUBSET 5					
29	S-15	-0.0838	-0.0918	0.0158	-0.1223
30	S-16	-0.1803	-0.1381	0.0634	-0.0104
31	S-31	-0.0028	-0.1058	0.2387	-0.0354
32	S-32	-0.0944	-0.0704	0.2407	-0.0074
33	S-33	-0.2298	-0.0673	0.0637	0.0201
34	S-34	-0.2805	-0.1038	0.0290	0.0657
35	S-35	-0.1643	0.0060	0.1473	-0.0420

DISCRIMINANT SCORES FOR SUBSETS

SUBSET 1	0.0323	0.1429	0.1245	0.0189
SUBSET 2	0.0814	0.1760	-0.0946	-0.0139
SUBSET 3	-0.2631	-0.0478	-0.1430	0.0157
SUBSET 4	0.1842	-0.1997	-0.0248	0.0040
SUBSET 5	-0.1480	-0.0816	0.1141	-0.0188
GRAND	0.0000	0.0000	0.0000	0.0000

CORRELATIONS BETWEEN INPUT VARIABLES AND DISCRIMINANT FUNCTIONS
INPUT VARIABLES IN ROWS, DISCRIMINANT FUNCTIONS IN COLUMNS

VARIABLE 1	0.6244	0.1820	-0.5112	-0.0997
VARIABLE 2	0.8573	-0.2012	-0.1606	0.0422
VARIABLE 3	-0.3206	0.4561	0.5229	0.1490
VARIABLE 4	0.7211	-0.2462	-0.4096	-0.0311

EUCLIDEAN DISTANCES FROM SUBSET MEANS (IN DISCRIMINANT SPACE)

		1	2	3	4	5	RANK
SUBSET 1							
1	S- 2	0.20	0.32	0.45	0.53	0.37	1
2	S- 3	0.14	0.24	0.44	0.48	0.36	1
3	S- 8	0.11	0.26	0.51	0.44	0.34	1
4	S- 9	0.24	0.38	0.59	0.50	0.38	1
5	S-10	0.05	0.23	0.45	0.41	0.30	1
6	S- 7	0.14	0.22	0.42	0.36	0.31	1
7	S-30	0.15	0.26	0.41	0.27	0.22	1
SUBSET 2							
8	S- 6	0.17	0.16	0.45	0.44	0.40	2
9	S- 4	0.26	0.19	0.49	0.53	0.48	2

10	S- 5	0.33	0.21	0.56	0.58	0.56	2
11	S-11	0.14	0.14	0.40	0.31	0.31	2
12	S-12	0.23	0.04	0.40	0.37	0.39	2
13	S-13	0.40	0.22	0.37	0.32	0.43	2
14	S-14	0.35	0.25	0.45	0.38	0.42	2
15	S-21	0.37	0.19	0.48	0.48	0.50	2
SUBSET 3							
16	S- 1	0.53	0.60	0.49	0.68	0.51	3
17	S-17	0.41	0.44	0.14	0.43	0.17	3
18	S-18	0.52	0.52	0.14	0.51	0.29	3
19	S-19	0.62	0.53	0.23	0.61	0.46	3
20	S-20	0.54	0.40	0.28	0.57	0.48	3
SUBSET 4							
21	S-22	0.63	0.50	0.58	0.48	0.61	4
22	S-23	0.50	0.37	0.53	0.26	0.55	4
23	S-24	0.90	0.83	0.85	0.67	0.90	4
24	S-25	0.45	0.49	0.70	0.26	0.52	4
25	S-26	0.44	0.54	0.64	0.27	0.41	4
26	S-27	0.44	0.53	0.55	0.20	0.35	4
27	S-28	0.40	0.49	0.50	0.21	0.29	4
28	S-29	0.31	0.37	0.35	0.19	0.19	5
SUBSET 5							
29	S-15	0.32	0.35	0.28	0.32	0.16	5
30	S-16	0.36	0.44	0.24	0.38	0.08	5
31	S-31	0.28	0.44	0.47	0.34	0.19	5
32	S-32	0.28	0.45	0.42	0.41	0.14	5
33	S-33	0.34	0.43	0.21	0.44	0.10	5
34	S-34	0.41	0.48	0.19	0.48	0.18	5
35	S-35	0.25	0.39	0.32	0.44	0.10	5

NO. OF CORE ITEMS = 31
RATIO OF NO. OF CORE ITEMS TO TOTAL NO. OF ITEMS = 0.886

ITERATION 3

MEANS OVER SUBSETS (IN INITIAL SPACE)

SUBSET 1	-0.0485	0.1426	0.0554	-0.1075
SUBSET 2	0.0570	0.1264	-0.1456	-0.0794
SUBSET 3	-0.1795	-0.1126	-0.1437	0.1632
SUBSET 4	0.2560	-0.1011	0.0948	0.0505
SUBSET 5	-0.1264	-0.0923	0.1040	0.0274
GRAND	0.0	0.0000	0.0000	0.0000

STANDARD DEVIATIONS OVER SUBSETS (IN INITIAL SPACE)

SUBSET 1	0.0535	0.0928	0.0379	0.1284
SUBSET 2	0.0499	0.1523	0.0797	0.0839
SUBSET 3	0.0575	0.2181	0.1913	0.1263
SUBSET 4	0.1292	0.1349	0.2264	0.2743
SUBSET 5	0.0940	0.0435	0.0840	0.0832
GRAND	0.1715	0.1715	0.1715	0.1715

B = BETWEEN-SUBSETS SCATTER MATRIX (IN INITIAL SPACE)

```
DEGREES OF FREEDOM =    4


                    C.7901  0.0225  0.1084 -C.0833
                    0.0225  0.4733 -C.1549 -C.3355
                    0.1084 -0.1549  0.4437 -0.0101
                   -0.0833 -C.3355 -C.C1C1  C.2884


W = WITHIN-SUBSETS SCATTER MATRIX (IN INITIAL SPACE)
DEGREES OF FREEDOM =  30


                    0.2099 -0.C225 -0.1C84  C.0833
                   -0.0225  0.5267  0.1549  C.3355
                   -0.1084  0.1549  0.5563  0.0101
                    0.0833  C.3355  C.01C1  C.7117


TRACE OF B =      1.9955 ;      TRACE OF W =      2.0046
(TRACE CF B)/(TRACE OF W) =      0.995464


EIGENVALUES CF  W-INVERSE * B

  5.33899  3.25915  C.61495  0.C0727

TRACE CF W-INVERSE * B =       9.22037

PERCENTAGE CF TRACE DUE TO EACH EIGENVALUE

   57.90     35.35     6.67     0.08

CUMULATIVE PERCENTAGE OF TRACE CUE TO EACH EIGENVALUE

   57.9C     93.25    99.92   100.0C


EIGENVECTCRS OF  W-INVERSE * B, NCRMALIZEC, AS COLUMNS


                    0.9052 -0.2865 -0.3123  C.0307
                    0.2623  C.7261  0.1549  0.6164
                    0.1518 -0.4321  C.8591  C.2286
                   -0.2978 -0.4516 -0.3748  0.7529


RATIO OF WITHIN TC TOTAL SCATTER IN 1,..., 4-DIMENSIONAL DISCRIMINANT SPACE

 C.16E 00 C.37E-C1 0.23E-C1 C.23E-C1

WILKS LAMBDA =   C.22769373E-01


TEST FCR HYPCTHESIS THAT SUBSET MEAN VECTORS ARE EQUAL :
F =     12.723 ;      DEGREES OF FREEDOM : N1 =   16,  N2 =       83
THE TEST ASSUMES NORMALITY AND A CCMMON DISPERSION MATRIX


DISCRIMINANT SCORES FOR ITEMS
```

SUBSET 1

1	S- 2	-0.0555	0.2092	0.1087	0.1861
2	S- 3	0.0022	0.2117	0.0711	0.1275
3	S- 8	0.0714	0.1720	0.1625	-0.0769
4	S- 9	0.0665	0.1563	0.2496	-0.1839
5	S-10	0.0396	0.1469	0.1246	-0.0287
6	S- 7	0.0546	0.0983	0.0570	0.1276
7	S-30	0.0583	0.0003	0.1023	-0.0249
SUBSET 2					
8	S- 6	0.0756	0.2091	0.0001	0.1061
9	S- 4	0.0604	0.2921	-0.0552	0.1344
10	S- 5	0.1132	0.3650	-0.0952	0.0751
11	S-11	0.0803	0.0910	0.0002	0.0377
12	S-12	0.0825	0.1490	-0.1029	0.0125
13	S-13	0.0920	0.0079	-0.2368	-0.0467
14	S-14	0.0939	0.0732	-0.0949	-0.2367
15	S-21	0.0928	0.2065	-0.1635	-0.1897
SUBSET 3					
16	S- 1	-0.2499	0.0636	0.0480	0.4537
17	S-17	-0.2440	-0.1149	-0.0311	-0.0240
18	S-18	-0.3106	-0.1521	-0.1203	-0.0570
19	S-19	-0.3239	-0.0690	-0.3020	-0.1359
20	S-20	-0.1838	0.0627	-0.3245	-0.1615
SUBSET 4					
21	S-22	0.1378	-0.1030	-0.2922	-0.3766
22	S-23	0.2550	-0.1101	-0.2511	0.0139
23	S-24	0.2734	-0.3290	-0.3559	0.5674
24	S-25	0.3541	-0.1597	0.1392	-0.0858
25	S-26	0.2034	-0.2335	0.2208	-0.1083
26	S-27	0.1284	-0.2889	0.1540	0.0011
27	S-28	0.0800	-0.2491	0.1533	0.0251
SUBSET 5					
28	S-29	0.0088	-0.1556	0.0488	0.0085
29	S-15	-0.0853	-0.0903	0.0141	-0.1226
30	S-16	-0.1839	-0.1343	0.0611	-0.0114
31	S-31	-0.0077	-0.1060	0.2383	-0.0363
32	S-32	-0.0986	-0.0686	0.2395	-0.0089
33	S-33	-0.2321	-0.0623	0.0610	0.0187
34	S-34	-0.2834	-0.0975	0.0260	0.0643
35	S-35	-0.1655	0.0093	0.1452	-0.0436

DISCRIMINANT SCORES FOR SUBSETS

SUBSET 1	0.0339	0.1421	0.1251	0.0181
SUBSET 2	0.0863	0.1742	-0.0935	-0.0134
SUBSET 3	-0.2625	-0.0420	-0.1460	0.0150
SUBSET 4	0.2046	-0.2105	-0.0331	0.0052
SUBSET 5	-0.1309	-0.0882	0.1043	-0.0164
GRAND	0.0000	0.0000	0.0000	0.0000

CORRELATIONS BETWEEN INPUT VARIABLES AND DISCRIMINANT FUNCTIONS
INPUT VARIABLES IN ROWS, DISCRIMINANT FUNCTIONS IN COLUMNS

VARIABLE 1	0.6345	0.1687	-0.5043	-0.0941
VARIABLE 2	0.8543	-0.2194	-0.1509	0.0483
VARIABLE 3	-0.3174	0.4628	0.5204	0.1434

VARIABLE 4 0.7203 -0.2615 -0.4017 -0.0244

EUCLIDEAN DISTANCES FROM SUBSET MEANS (IN DISCRIMINANT SPACE)

		1	2	3	4	5	RANK
SUBSET 1							
1	S- 2	0.20	0.32	0.45	0.54	0.37	1
2	S- 3	0.14	0.24	0.44	0.49	0.36	1
3	S- 8	0.11	0.26	0.51	0.46	0.34	1
4	S- 9	0.24	0.38	0.59	0.52	0.38	1
5	S-10	0.05	0.23	0.45	0.43	0.29	1
6	S- 7	0.14	0.22	0.42	0.38	0.30	1
7	S-30	0.15	0.26	0.41	0.29	0.21	1
SUBSET 2							
8	S- 6	0.17	0.16	0.45	0.45	0.40	2
9	S- 4	0.26	0.19	0.49	0.54	0.48	2
10	S- 5	0.33	0.21	0.56	0.59	0.56	2
11	S-11	C.14	0.14	0.40	0.33	0.30	2
12	S-12	0.23	0.04	0.40	0.39	0.38	2
13	S-13	0.40	0.22	0.37	0.32	C.42	2
14	S-14	0.35	0.25	0.45	0.39	0.41	2
15	S-21	0.37	0.19	0.48	0.49	0.49	2
SUBSET 3							
16	S- 1	0.53	0.60	0.49	0.70	0.51	3
17	S-17	0.41	0.44	0.14	0.46	0.18	3
18	S-18	0.52	0.52	0.14	0.53	0.30	3
19	S-19	0.62	0.53	0.23	0.63	0.47	3
20	S-20	0.54	0.40	0.28	0.58	0.48	3
SUBSET 4							
21	S-22	0.63	0.50	0.58	0.48	0.60	4
22	S-23	0.50	0.37	0.53	0.25	0.53	4
23	S-24	0.90	0.83	0.85	0.66	0.88	4
24	S-25	0.45	0.49	0.70	0.25	0.50	4
25	S-26	0.44	0.54	0.64	0.28	0.39	4
26	S-27	0.44	0.53	0.55	0.22	0.33	4
27	S-28	0.40	0.49	0.50	0.23	0.27	4
SUBSET 5							
28	S-29	0.31	0.37	0.35	0.22	0.17	5
29	S-15	0.32	0.35	0.28	0.34	C.15	5
30	S-16	0.36	0.44	0.24	0.41	0.08	5
31	S-31	0.28	0.44	0.47	0.36	0.18	5
32	S-32	0.28	0.45	0.42	0.43	0.14	5
33	S-33	C.34	0.43	0.21	0.47	C.12	5
34	S-34	0.41	0.48	0.19	0.51	0.19	3
35	S-35	0.25	0.39	0.32	0.47	C.11	5

NO. OF CORE ITEMS = 30
RATIO OF NO. OF CORE ITEMS TO TOTAL NO. OF ITEMS = 0.857

ITERATION 4

MEANS OVER SUBSETS (IN INITIAL SPACE)

SUBSET 1	-0.0485	0.1426	0.0554	-0.1075
SUBSET 2	0.0570	C.1264	-0.1456	-C.C794

SUBSET 3	-0.1887	-0.1108	-0.1137	0.1638
SUBSET 4	0.2560	-0.1011	0.0948	0.0505
SUBSET 5	-0.1109	-0.0910	0.1137	0.0074
GRAND	0.0000	0.0000	0.0000	0.0000

STANDARD DEVIATIONS OVER SUBSETS (IN INITIAL SPACE)

SUBSET 1	0.0535	0.0928	0.0379	0.1284
SUBSET 2	0.0499	0.1523	0.0797	0.0839
SUBSET 3	0.0562	0.1952	0.1862	0.1130
SUBSET 4	0.1292	0.1349	0.2264	0.2743
SUBSET 5	0.0898	0.0468	0.0857	0.0660
GRAND	0.1715	0.1715	0.1715	0.1715

B = BETWEEN-SUBSETS SCATTER MATRIX (IN INITIAL SPACE)
DEGREES OF FREEDOM = 4

0.8010	0.0242	0.1251	-0.1004
0.0242	0.4733	-0.1558	-0.3370
0.1251	-0.1558	0.4220	-0.0216
-0.1004	-0.3370	-0.0216	0.3107

W = WITHIN-SUBSETS SCATTER MATRIX (IN INITIAL SPACE)
DEGREES OF FREEDOM = 30

0.1990	-0.0242	-0.1251	0.1004
-0.0242	0.5267	0.1558	0.3370
-0.1251	0.1558	0.5780	0.0216
0.1004	0.3370	0.0216	0.6894

TRACE OF B = 2.0070 ; TRACE OF W = 1.9931
(TRACE OF B)/(TRACE OF W) = 1.006968

EIGENVALUES OF W-INVERSE * B

6.64543 3.21804 0.56974 0.01202

TRACE OF W-INVERSE * B = 10.44523

PERCENTAGE OF TRACE DUE TO EACH EIGENVALUE

63.62 30.81 5.45 0.12

CUMULATIVE PERCENTAGE OF TRACE DUE TO EACH EIGENVALUE

63.62 94.43 99.88 100.00

EIGENVECTORS OF W-INVERSE * B, NORMALIZED, AS COLUMNS

0.8911	-0.3017	-0.3374	0.0322
0.2639	0.7278	0.1056	0.6242

0.1730	-0.4145	C.8530	C.2659
-0.3261	-0.4555	-C.3840	C.7339

RATIO OF WITHIN TO TOTAL SCATTER IN 1,..., 4-DIMENSIONAL DISCRIMINANT SPACE

0.13E 00 C.31E-C1 C.20E-C1 0.20E-01

WILKS LAMBDA = 0.19519631E-01

TEST FOR HYPOTHESIS THAT SUBSET MEAN VECTORS ARE EQUAL :
F = 13.649 ; DEGREES OF FREEDOM : N1 = 16, N2 = 83
THE TEST ASSUMES NORMALITY AND A COMMON DISPERSION MATRIX

DISCRIMINANT SCORES FOR ITEMS

SUBSET 1					
1	S- 2	-0.0529	0.2123	C.0985	C.1889
2	S- 3	0.0044	0.2133	0.0613	C.1297
3	S- 8	0.0795	0.1747	0.1592	-0.0697
4	S- 9	0.0791	0.1607	C.2508	-0.1732
5	S-10	0.0455	0.1491	0.1210	-0.0235
6	S- 7	0.0553	0.0992	0.0479	0.1304
7	S-30	0.0621	0.0019	0.1012	-C.0200
SUBSET 2					
8	S- 6	0.0757	0.2085	-0.0112	0.1066
9	S- 4	0.0589	C.2904	-0.0689	C.1323
10	S- 5	0.1120	0.3619	-0.1098	0.0721
11	S-11	C.0804	C.0903	-C.006C	0.0386
12	S-12	0.0801	0.1459	-C.1093	C.0092
13	S-13	0.0848	0.0018	-0.2378	-0.0548
14	S-14	0.0951	0.0698	-0.0902	-0.2391
15	S-21	0.0921	0.2017	-0.1634	-C.1952
SUBSET 3					
16	S- 1	-0.2548	C.0677	0.0371	C.4514
17	S-17	-0.2455	-0.1133	-0.0194	-C.0286
18	S-18	-0.3149	-0.1518	-0.1040	-0.0660
19	S-19	-0.3322	-0.0727	-0.2837	-C.1524
20	S-20	-0.1914	0.0572	-0.3127	-0.1771
21	S-34	-0.2842	-0.0942	0.0351	0.0614
SUBSET 4					
22	S-22	0.1330	-0.1112	-0.2793	-0.3857
23	S-23	0.2452	-C.1179	-0.2574	C.0079
24	S-24	0.2493	-0.3382	-C.3799	C.5574
25	S-25	0.3585	-C.1601	0.1339	-0.0746
26	S-26	0.2103	-0.2308	0.2229	-C.0959
27	S-27	0.1309	-0.2867	0.1557	0.0098
28	S-28	0.0824	-0.2465	0.1548	0.0330
SUBSET 5					
29	S-29	0.0089	-0.1546	0.0515	C.0109
30	S-15	-0.0836	-0.0893	0.0238	-0.1230
31	S-16	-0.1827	-0.1312	C.0706	-C.0113
32	S-31	-0.0000	-0.1008	0.2419	-0.0266
33	S-32	-C.0910	-0.0625	0.2443	-C.0006
34	S-33	-0.2307	-0.0587	0.0694	C.0179
35	S-35	-0.1598	0.0140	0.1521	-0.0402

DISCRIMINANT SCORES FOR SUBSETS

SUBSET 1	0.0390	0.1445	0.1200	0.0233
SUBSET 2	0.0849	0.1713	-0.0996	-0.0163
SUBSET 3	-0.2705	-0.0512	-0.1079	0.0148
SUBSET 4	0.2014	-0.2131	-0.0356	0.0074
SUBSET 5	-0.1056	-0.0833	0.1220	-0.0247
GRAND	0.0000	0.0000	0.0000	0.0000

CORRELATIONS BETWEEN INPUT VARIABLES AND DISCRIMINANT FUNCTIONS
INPUT VARIABLES IN ROWS, DISCRIMINANT FUNCTIONS IN COLUMNS

VARIABLE 1	0.3588	0.2010	-0.4328	-0.1931
VARIABLE 2	0.6485	-0.1787	-0.2424	-0.0044
VARIABLE 3	-0.1108	0.4332	0.5105	0.1413
VARIABLE 4	0.4607	-0.2165	-0.4120	-0.1661

EUCLIDEAN DISTANCES FROM SUBSET MEANS (IN DISCRIMINANT SPACE)

	1	2	3	4	5	RANK
SUBSET 1						
1 S- 2	0.20	0.32	0.44	0.54	0.37	1
2 S- 3	0.14	0.24	0.43	0.49	0.36	1
3 S- 8	0.11	0.26	0.50	0.46	0.32	1
4 S- 9	0.24	0.38	0.58	0.52	0.36	1
5 S-10	0.05	0.23	0.44	0.43	0.28	1
6 S- 7	0.14	0.22	0.41	0.38	0.30	1
7 S-30	0.15	0.26	0.40	0.29	0.19	1
SUBSET 2						
8 S- 6	0.17	0.16	0.45	0.45	0.39	2
9 S- 4	0.26	0.19	0.49	0.54	0.48	2
10 S- 5	0.33	0.21	0.57	0.59	0.56	2
11 S-11	0.14	0.14	0.39	0.33	0.29	2
12 S-12	0.23	0.04	0.40	0.39	0.38	2
13 S-13	0.40	0.22	0.39	0.32	0.42	2
14 S-14	0.35	0.25	0.46	0.39	0.39	2
15 S-21	0.37	0.19	0.49	0.49	0.48	2
SUBSET 3						
16 S- 1	0.53	0.60	0.48	0.70	0.53	3
17 S-17	0.41	0.44	0.12	0.46	0.20	3
18 S-18	0.52	0.52	0.14	0.53	0.32	3
19 S-19	0.62	0.53	0.25	0.63	0.48	3
20 S-20	0.54	0.40	0.31	0.58	0.49	3
21 S-34	0.41	0.48	0.16	0.51	0.22	3
SUBSET 4						
22 S-22	0.63	0.50	0.60	0.48	0.59	4
23 S-23	0.50	0.37	0.54	0.25	0.52	4
24 S-24	0.90	0.83	0.85	0.66	0.88	4
25 S-25	0.45	0.49	0.69	0.25	0.47	4
26 S-26	0.44	0.54	0.62	0.28	0.37	4
27 S-27	0.44	0.53	0.53	0.22	0.32	4
28 S-28	0.40	0.49	0.48	0.23	0.26	4
SUBSET 5						
29 S-29	0.31	0.37	0.34	0.22	0.16	5
30 S-15	0.32	0.35	0.27	0.34	0.14	5
31 S-16	0.36	0.44	0.22	0.41	0.11	5

32	S-31	0.28	0.44	0.45	0.36	0.16	5
33	S-32	0.28	0.45	0.40	0.43	0.13	5
34	S-33	0.34	0.43	0.18	0.47	0.14	5
35	S-35	0.25	0.39	0.30	0.47	0.12	5

NO. OF CORE ITEMS = 30
RATIO OF NO. OF CORE ITEMS TO TOTAL NO. OF ITEMS = 0.857

SAMPLE PROBLEM USING HYPOTHETICAL DATA

SUMMARY OF 4 ITERATIONS

		1	2	3	4
SUBSET 1					
1	S- 2	1	1	1	1
2	S- 3	1	1	1	1
3	S- 6	2	2	2	2
4	S- 8	1	1	1	1
5	S- 9	1	1	1	1
6	S-10	1	1	1	1
SUBSET 2					
7	S- 4	2	2	2	2
8	S- 5	2	2	2	2
9	S- 7	1	1	1	1
10	S-11	2	2	2	2
11	S-12	2	2	2	2
12	S-13	2	2	2	2
13	S-14	2	2	2	2
14	S-21	2	2	2	2
SUBSET 3					
15	S- 1	3	3	3	3
16	S-17	3	3	3	3
17	S-18	3	3	3	3
18	S-19	3	3	3	3
19	S-20	3	3	3	3
SUBSET 4					
20	S-22	4	4	4	4
21	S-23	4	4	4	4
22	S-24	4	4	4	4
23	S-25	4	4	4	4
24	S-26	4	4	4	4
25	S-27	4	4	4	4
26	S-28	4	4	4	4
27	S-29	4	5	5	5
28	S-30	1	1	1	1
SUBSET 5					
29	S-15	5	5	5	5
30	S-16	5	5	5	5
31	S-31	5	5	5	5
32	S-32	5	5	5	5
33	S-33	5	5	5	5
34	S-34	5	5	3	3
35	S-35	5	5	5	5

Table 2.—Listing of input to hypothetical problem.

```
    1
SAMPLE PROBLEM USING HYPOTHETICAL DATA
   35    4    5    5            1    1    1    1                1    4       100.0
  6  8  5  9  7
(A4,A2,6X,4F6.0)
S- 2        000101000092000142000049
S- 3        000119000098000131000068
S- 6        000147000106000118000102
S- 8        000113000107000116000066
S- 9        000094000107000115000044
S-10        000111000104000117000069
END 1 ST SUBSET
S- 4        000157000101000124000092
S- 5        000178000104000119000097
S- 7        000128000108000116000109
S-11        000139000110000104000117
S-12        000157000107000100000118
S-13        000169000111000075000157
S-14        000145000109000079000107
S-21        000164000104000088000097
END 2 ND SUBSET
S- 1        000063000075000159000041
S-17        000048000077000111000032
S-18        000041000069000106000022
S-19        000066000062000097000017
S-20        000111000074000092000045
END 3 RD SUBSET
S-22        000170000117000039000164
S-23        000208000135000053000246
S-24        000237000148000058000366
S-25        000169000152000061000230
S-26        000114000137000073000175
S-27        000106000130000077000178
S-28        000097000123000086000156
S-29        000099000110000092000125
S-30        000111000111000102000105
END 4 TH SUBSET
S-15        000079000095000096000069
S-16        000049000086000111000047
S-31        000068000108000108000081
S-32        000048000096000121000044
S-33        000042000078000123000020
S-34        000034000073000125000017
S-35        000048000084000125000014
END 5 TH SUBSET, END ALL DATA SET
```

REFERENCES

Anderson, T.W., 1958, An introduction to multivariate statistical analysis: John Wiley and Sons, Inc., New York, 374 p.

Anderson, T.W., 1963, Asymptotic theory for principal component analysis: Ann. Math. Statistics, v. 34, p. 122-148.

Ball, G.H., 1965, Data analysis in the social sciences: What about the details?, in 1965 Fall Joint Conf. Proc., Am. Federation Info. Processing Soc., p. 533-559.

Casetti, E., 1964, Classificatory and regional analysis by discriminant iterations: Office of Naval Research, Geography Branch, Tech. Rept. No. 12, ONR Task No. 389-135, 95 p.

Dixon, W.J., ed., 1967, BMD biomedical computer programs: Univ. California Pub. in Automatic Computation, no. 2, 600 p.

Eades, D.C., 1965, The inappropriateness of the correlation coefficient as a measure of taxonomic resemblance: Systematic Zoology, v. 14, p. 98-100.

Forgy, Edward, 1965 (abs.), Cluster analysis of multivariate data: Efficiency vs. interpretability of classifications: Biometrics, v. 21, p. 768.

Fortier, J.J., and Solomon, H., 1966, Clustering procedures, in Krishnaiah, P.R., ed., Multivariate analysis: Academic Press, Inc., New York, p. 493-506.

Friedman, H.P., and Rubin, J., 1967, On some invariant criteria for grouping data: Am. Stat. Assoc. Jour., v. 62, p. 1159-1178.

Goodall, D.W., 1966a, Numerical taxonomy of bacteria - some published data re-examined: Jour. General Microbiology, v. 42, p. 25-37.

Goodall, D.W., 1966b, A new similarity index based on probability: Biometrics, v. 22, p. 882-907.

Gower, J.C., 1967a, Multivariate analysis and multidimensional geometry: Statistician, v. 17, p. 13-28.

Gower, J.C., 1967b, A comparison of some methods of cluster analysis: Biometrics, v. 23, p. 623-637.

Hotelling, Harold, 1951, A generalized T test and measure of multivariate dispersion, in Second Berkeley Symposium on Mathematical Statistics and Probability Proc., p. 23-41.

Johnson, S.C., 1967, Hierarchical clustering schemes: Physchometrika, v. 32, p. 241-254.

MacQueen, J., 1966, Some methods for classification and analysis of multivariate observations, in Fifth Berkeley Symposium on Mathematical Statistics and Probability Proc., p. 281-297.

Minkoff, E.C., 1965, The effects on classification of slight alterations in numerical technique: Systematic Zoology, v. 14, p. 196-213.

Rao, C.R., 1952, Advanced statistical methods in biometric research: John Wiley and Sons, Inc., New York, 390 p.

Rohlf, F.J., and Sokal, R.R., 1965, Coefficients of correlation and distance in numerical taxonomy: Univ. Kansas Sci. Bull., v. 45, p. 3-27.

Sokal, R.R., and Rohlf, F.J., 1962, The comparison of dendrograms by objective methods: Taxon, v. 11, p. 33-40.

Sokal, R.R., and Sneath, P.H.A., 1963, Principles of numerical taxonomy: W.H. Freeman Co., San Francisco, 359 p.

Wilks, S.S., 1932, Certain generalizations in the analysis of variance: Biometrika, v. 24, p. 471-494.

Wilks, S.S., 1960, Multidimensional statistical scatter, in Olkin, I., et al, eds., Contributions to probability and statistics - essays in honor of Harold Hotelling: Stanford Univ. Press, p. 486-503.

Wilks, S.S., 1962, Mathematical statistics: John Wiley and Sons, Inc., New York, 644 p.

Williams, W.T., and Dale, M.B., 1965, Fundamental problems in numerical taxonomy: Adv. Botanical Research, v. 2, p. 35-68.

Wolleben, J.A., Pauken, R. J., and Dearien, J.A., 1968, FORTRAN IV program for multivariate paleontologic analysis using an IBM System/360 Model 40 computer, in Merriam, D.F., ed., Computer programs for multivariate analysis in geology: Kansas Geol. Survey Computer Contr. 20, p. 1-12.

APPENDIX

Four theorems that have a bearing on computational options provided in the program are stated below. Proofs are omitted and will be given, along with discussion of other aspects, in a future paper. In the interim, the interested reader may wish to prove these theorems for himself. Mimeographed proofs are also available from the writer upon request. For proof of item (i), Theorem (3), see also Friedman and Rubin (1967).

Theorem 1: Let $\underline{E}^{(1)}$ (nxp) be a score matrix in some $e^{(1)}$-space, and $\underline{E}^{(2)}$ (nxp) $= \underline{E}^{(1)}\underline{D}^{-1}(\sqrt{k_i})$ a nonsingular diagonal transform of $E^{(1)}$, where k_i ($i = 1, \ldots, p$) are some positive values. Let $\underline{C}^{(1)}$(pxp) and $\underline{C}^{(2)}$ (pxp) be covariance matrices, both assumed to be positive definite, in $e^{(1)}$- and $e^{(2)}$-spaces, respectively, Then, if $\underline{X}^{(1)}$ (nxp) and $\underline{X}^{(2)}$ (nxp) represent, respectively, orthonormalized forms of $\underline{E}^{(1)}$ and $\underline{E}^{(2)}$ under covariance option, $X^{(1)}$ and $X^{(2)}$ are related orthogonally (it is implicitly assumed that all principal components are retained in both cases). More specifically, if we let λ_i be the i-th eigenvalue of $\underline{C}^{(1)}$, β_i the i-th eigenvalue of $\underline{C}^{(2)}$, and if columns of $\underline{P}$ (pxp) and $\underline{Q}$ (pxp) contain, respectively, normalized eigenvectors of $\underline{C}^{(1)}$ and $\underline{C}^{(2)}$, then $\underline{X}^{(1)} = \underline{X}^{(2)}\,\underline{A}$, where

$$\underline{A} = \underline{D}(\sqrt{\beta_i})\ \underline{Q}'\ \underline{D}(\sqrt{k_i})\ \underline{P}\ \underline{D}^{-1}(\sqrt{\lambda_i})$$

is orthogonal. In the special instance where $e^{(1)}$-variables are uncorrelated, or when $k_i = k$ for all i, then $\underline{A} = \underline{I}$, i.e., $X^{(1)} = X^{(2)}$. If orthonormalization is performed under the correlation instead of the covariance option, the respective orthonormalized score matrices obtained from $e^{(1)}$- and $e^{(2)}$-spaces are always equal, i.e., using the notation analogous to that given above, always $\underline{A} = \underline{I}$ and $X^{(1)} = X^{(2)}$.

Theorem 2: Let $\underline{E}$ (nxp) be a score matrix in some e-space, and $\underline{C}$ (pxp) and $\underline{R}$ (pxp) positive definite covariance and correlation matrices, respectively, in this space. Then, if $\underline{X}^{(c)}$ (nxp) and $\underline{X}^{(r)}$ (nxp) represent orthonormalized forms of $\underline{E}$ under covariance and correlation options, respectively, $\underline{X}^{(c)}$ and $\underline{X}^{(r)}$ are orthogonally related (it is implicitly assumed that all principal components are retained in both cases). More specifically, if we let s_i be the standard deviation of e_i, λ_i the i-th eigenvalue of $\underline{C}$, β_i the i-th eigenvalue of $\underline{R}$, and if columns of $\underline{P}$ (pxp) and $\underline{Q}$ (pxp) contain, respectively, normalized eigenvectors of $\underline{C}$ and $\underline{R}$, then $\underline{X}^{(c)} = \underline{X}^{(r)}\,\underline{A}$, where

$$\underline{A} = \underline{D}(\sqrt{\beta_i})\ \underline{Q}'\ \underline{D}(s_i)\ \underline{P}\ \underline{D}^{-1}(\sqrt{\lambda_i})$$

is orthogonal. In the special situation where e-variables are uncorrelated, $\underline{A} = \underline{I}$, i.e., $\underline{X}^{(c)} = \underline{X}^{(r)}$.

Theorem 3: Let $\overline{X}$ (nxp) be a partitioned score matrix in some x-space, $\overline{B}$ (pxp) and $\underline{W}$ (pxp), the latter assumed to be positive definite, the between and within scatter matrices in this space, and $\underline{Y}$ (nxp) the discriminant score matrix derived from $\underline{X}$ under the assumption that eigenvectors associated with discriminant functions (i.e., eigenvectors of $\underline{W}^{-1}\underline{B}$) are left nonnormalized. Then

(i) The ordinary Euclidean distances in the discriminant space, assuming all discriminant functions are used, are identical to corresponding Mahalanobis distances in the initial space; that is, in our notation, if we consider squared distances between, say, the k-th and l-th items in the h-th group,

$$(\underline{y}_{hk.} - \underline{y}_{hl.})(\underline{y}_{hk.} - \underline{y}_{hl.})' = (\underline{x}_{hk.} - \underline{x}_{hl.})\ \underline{W}^{-1}\ (\underline{x}_{hk.} - \underline{x}_{hl.})'.$$

(ii) If we let $\underline{T}_{(y)}$ (pxp) be the total scatter matrix in the discriminant space, tr $\underline{T}_{(y)} = p + \text{tr}\,\underline{W}^{-1}\underline{B}$.

(iii) Discriminant functions are uncorrelated, with the i-th discriminant function having a sample variance $(1 + \theta_i) / (n - 1)$, where θ_i is the i-th eigenvalue of $\underline{W}^{-1}\underline{B}$.

Theorem 4: Let $\underline{X}$ (nxp) be a partitioned score matrix in some orthonormal x-space, $\underline{B}$ (pxp) and $\underline{W}$ (pxp), the latter assumed to be positive definite, the between and within scatter matrices in this space, and Y^* (nxp) the discriminant score matrix derived from $\overline{X}$ under the assumption that eigenvectors associated with discriminant functions (i.e., eigenvectors of $\underline{W}^{-1}\underline{B}$) are normalized. Then

(i) $\underline{Y}^*$ represents an orthogonal transformation of $\underline{X}$.

(ii) Discriminant functions are uncorrelated each with constant sample variance $1/(n-1)$.

(iii) If we let $\underline{V}^*$ (pxp) be the matrix whose columns are normalized eigenvector of $\underline{W}^{-1}\underline{B}$, the element v^*_{ij} of $\underline{V}^*$ represents the correlation between the i-th orthonormal variable x_i and the j-th discriminant function y^*_j.

KANSAS GEOLOGICAL SURVEY COMPUTER PROGRAM
THE UNIVERSITY OF KANSAS, LAWRENCE

PROGRAM ABSTRACT

Title (If subroutine state in title):

Multivariate procedures and FORTRAN IV program for evaluation and improvement of classifications.

Date: November, 1968

Author, organization: Ferruh Demirmen, Geology Department,
Stanford University

Direct inquiries to: Author, or

Name: D. F. Merriam

Address: Kansas Geological Survey
Lawrence, Kansas 66044

Purpose/description: Evaluates classifications by three criteria that measure the degree of "compactness" of a partition, and improves classifications by the nearest neighbor algorithm in discriminant space. Also performs principal component analysis, linear discriminant analysis, and one-way multivariate analysis of variance.

Mathematical method: Described in text. Matrix inversion and direct computation of determinants avoided.

Restrictions, range: Up to 300 items, 25 groups, 30 input variables, and 25 iterations allowed for each classification. More than one classification can be processed in one run. Storage requirements can be readjusted easily.

Computer manufacturer: IBM

Model: System/360, Model 67

Programming language: FORTRAN IV, Level H

Memory required: 345 K

Approximate running time:

Special peripheral equipment required: None

Remarks (special compilers or operating systems, required word lengths, number of successful runs, other machine versions, additional information useful for operation or modification of program)

Subroutines CORRE, ARRAY, and EIGEN, provided in the IBM Scientific Subroutine Package, must be available in the system. All options were tested repeatedly and successfully at Stanford University. Compilation time on IBM 360/67 is 25-30 seconds.

COMPUTER CONTRIBUTIONS

1. Mathematical simulation of marine sedimentation with IBM 7090/7094 computers, by J.W. Harbaugh, 1966 . . . . $1.00
2. A generalized two-dimensional regression procedure, by J.R.Dempsey, 1966. . . . . . . $0.50
3. FORTRAN IV and MAP program for computation and plotting of trend surfaces for degrees 1 through 6, by Mont O'Leary, R.H. Lippert, and O.T. Spitz, 1966 . . . . $0.75
4. FORTRAN II program for multivariate discriminant analysis using an IBM 1620 computer, by J.C.Davis and R.J. Sampson, 1966 . . . . $0.50
5. FORTRAN IV program using double Fourier series for surface fitting of irregularly spaced data, by W.R. James, 1966. . . . . $0.75
6. FORTRAN IV program for estimation of cladistic relationships using the IBM 7040, by R.L. Bartcher, 1966 . . . . $1.00
7. Computer applications in the earth sciences: Colloquium on classification procedures, edited by D.F. Merriam, 1966 . . . . $1.00
8. Prediction of the performance of a solution gas drive reservoir by Muskat's Equation, by Apolonio Baca, 1967 . . . . $1.00
9. FORTRAN IV program for mathematical simulation of marine sedimentation with IBM 7040 or 7094 computers, by J.W. Harbaugh and W.J. Wahlstedt, 1967 . . . . $1.00
10. Three-dimensional response surface program in FORTRAN II for the IBM 1620 computer, by R.J.Sampson and J.C.Davis, 1967. . . . . $0.75
11. FORTRAN IV program for vector trend analyses of directional data, by W.T. Fox, 1967 . . $1.00
12. Computer applications in the earth sciences: Colloquium on trend analysis, edited by D.F. Merriam and N.C. Cocke, 1967 . . . . $1.00
13. FORTRAN IV computer programs for Markov chain experiments in geology, by W.C.Krumbein, 1967 . . . . $1.00
14. FORTRAN IV programs to determine surface roughness in topography for the CDC 3400 computer, by R. D. Hobson, 1967 . . . . $1.00
15. FORTRAN II program for progressive linear fit of surfaces on a quadratic base using an IBM 1620 computer, by A.J. Cole, C. Jordan, and D. F. Merriam, 1967 . . . . $1.00
16. FORTRAN IV program for the GE 625 to compute the power spectrum of geological surfaces, by J.E. Esler and F.W. Preston, 1967. . . . . $0.75
17. FORTRAN IV program for Q-mode cluster analysis of nonquantitative data using IBM 7090/7094 computers, by G.F. Bonham-Carter, 1967 . . . . $1.00
18. Computer applications in the earth sciences: Colloquium on time-series analysis, D. F. Merriam, editor, 1967. . . . . $1.00
19. FORTRAN II time-trend package for the IBM 1620 computer, by J.C. Davis and R.J. Sampson, 1967 . . . . $1.00
20. Computer programs for multivariate analysis in geology, D.F. Merriam, editor, 1968. . . $1.00
21. FORTRAN IV program for computation and display of principal components, by W.J. Wahlstedt and J.C. Davis, 1968. . . . . $1.00
22. Computer applications in the earth sciences: Colloquium on simulation, D.F. Merriam and N.C. Cocke, editors, 1968 . . . . $1.00
23. Computer programs for automatic contouring, by D.B. McIntyre, D.D. Pollard, and R. Smith, 1968. . . . . $1.50
24. Mathematical model and FORTRAN IV program for computer simulation of deltaic sedimentation, by G.F. Bonham-Carter and A.J. Sutherland, 1968 . . . . $1.00
25. FORTRAN IV CDC 6400 computer program for analysis of subsurface fold geometry, by E.H.T. Whitten, 1968 . . . . $1.00
26. FORTRAN IV computer program for simulation of transgression and regression with continuous-time Markov models, by W.C. Krumbein, 1968. . . . . $1.00
27. Stepwise regression and nonpolynomial models in trend analysis, by A.T. Miesch and J.J. Connor, 1968. . . . . $1.00
28. KWIKR8 a FORTRAN IV program for multiple regression and geologic trend analysis, by J.E. Esler, P.F. Smith, and J.C.Davis, 1968. . . . . $1.00
29. FORTRAN IV program for harmonic trend analysis using double Fourier series and regularly gridded data for the GE 625 computer, by J.W. Harbaugh and M.J. Sackin, 1968. . . $1.00
30. Sampling a geological population (workshop on experiment in sampling), by J.C. Griffiths and C. W. Ondrick, 1968. . . . . $1.00
31. Multivariate procedures and FORTRAN IV program for evaluation and improvement of classifications, by Ferruh Demirmen, 1969. . . . . $1.00