A MULTIPLE SEPARATION FUNCTION FOR PATTERN CLASSIFICATION

Kikumi Tatsuoka

REPORT R-313 AUGUST, 1966
This work was supported in part by the Joint Services Electronics Program (U.S. Army, U.S. Navy, and U.S. Air Force) under Contract DA 28 043 AMC 00073(E); and in part by the Office of Education OE-6-10-184.

Reproduction in whole or in part is permitted for any purpose of the United States Government.

Distribution of this report is unlimited. Qualified requesters may obtain copies of this report from DDC.
ABSTRACT

In this paper we develop a method for classifying any object which can be represented as a point in n-space into one of m given subsets by constructing a linear function which is derived from the support function in linear topological space, an approach hitherto unexplored in this connection.

Greenberg and Konheim (1964) discussed this problem of classification scheme by presenting two possible procedures: one is to use pairwise class separation, and the other is to construct m linear functions which separate one of the sets from all the others. The former requires the construction of m(m-1)/2 linear functions, and the latter places more stringent restrictions on the m subsets than the former. We show, in this report, how to construct m instead of m(m-1)/2, linear functions under the more liberal assumptions of the first procedure, thus combining the merits of both procedures.

The application of our procedure to a problem in the classification of sequences of student responses on the PLATO teaching system is discussed. In that connection, a possible further development, an even simpler method requiring only one linear function, was explored. The validity of this simpler method could, however, be established only under rather restrictive conditions.
ACKNOWLEDGMENT

I gratefully acknowledge the helpful advice and encouragement given by Dr. Sylvian R. Ray in discussing an earlier version of this paper. Dr. John A. Easley and my husband Maurice Tatsuoka made many helpful suggestions concerning expository style in preparing the current draft.
0. INTRODUCTION

Greenberg and Konheim [1] discuss a technique for pattern classification of alphabetic characters described by the output of a transducer -- a computer program that converts the sensed arrays of light and dark signals derived from a character to a vector in n-dimensional space. The output of their processor program, simply operated, was the identification of the class to which the input (unknown) pattern belonged. However, they also employed their program iteratively to determine the boundaries which best separated representative patterns used for "training" the pattern classification system. We have in mind a more general class of uses which includes the classification of sequences of events ("dope records") obtained while students are interacting with the PLATO instructional system. This use constitutes a case of fact retrieval, since information (how many, who, when, in what order, etc.) can be requested regarding response-sequence categories, specified in advance either by their boundaries or by giving examples. For many uses it is anticipated that a simpler procedure which is broader in scope of applicability can be employed and will have considerable practical value for computational purposes.

More precisely, denote by Ω the space of patterns, in which it is assumed that m mutually disjoiing patterns, Ω₁, ..., Ωₘ have been identified a priori corresponding to the distinct responses required of the classification system. Let (α₁, α₂, ..., αₙ) be the vector (set of measurements produced by the transducer) corresponding to a given pattern and Rⁿ be a linear vector space whose base is (e₁, e₂, ..., eₙ), the usual standard orthogonal coordinate system.
By the operation of the transducer, each \( \Omega_i \) will be put into correspondence with a subset of \( \mathbb{R}^n \), say \( A_1 \ (i = 1, \ldots, m) \) and \( A_1, A_2, \ldots, A_m \) will be mutually disjoint. The processor must then be able to classify any \( \omega \) in \( \Omega \) by operating the output \( x(\omega) \) of the transducer "measurement" of \( \omega \).

Linear separability in pattern classification refers to the feasibility of classifying objects in \( \Omega \) by constructing linear boundaries between the sets \( A_1, \ldots, A_m \). Greenberg and Konheim [1] mentioned two possibilities for solving this problem. One of these calls for the construction of \( m \) linear functions \( L_1, L_2, \ldots, L_m \) which are to separate the sets \( A_1 \) and \( \bigcup_{i \neq j} A_j \), provided \( A_1 \) and \( \bigcup_{i \neq j} A_j \) are convex disjoint finite sets. This assumption becomes too stringent for practical applications because in applying the construction method of linear functions they must assume that the complement of every subset in the set of categories is itself convex.

In Section 1, the mathematical foundations of our method are presented. Most of the proofs are omitted, for they may be found in any treatise on linear operators, Banach space, and Hilbert space.

In Section 2 of this paper we introduce and discuss a method of constructing a linear separation function in the case when only two sets are given, by using the support function, which is different from the method Greenberg and Konheim constructed which used a training set to establish boundaries.

In Section 3, we develop our method of constructing a linear separation function in the case when \( m \) convex disjoint sets are given.
In Section 4, we discuss the application of our method to the classification of sequences of student responses made on the PLATO instructional system. (See Bitzer, Lyman, and Easley [4].)

1. MATHEMATICAL FOUNDATIONS

Let \( \mathbb{R}^n \) be an \( n \)-dimensional Euclidean space.

**Definition:** A set \( K \subseteq \mathbb{R}^n \) is convex if and only if \( x, y, \in K \) and \( 0 \leq a \leq 1 \) imply \( ax + (1-a)y \in K \).

**Property 1:** The intersection of an arbitrary family of convex subsets of the linear space \( \mathbb{R}^n \) is convex.

**Property 2:** Let \( x_1, x_2, \ldots, x_n \) be points in the convex set \( K \) and let \( a_1, a_2, \ldots, a_n \) be non-negative numbers such that \( \sum a_i = 1 \). Then, \( \sum a_i x_i \in K \).

**Lemma 1:** If \( K_1 \) and \( K_2 \) are arbitrary convex subsets of \( \mathbb{R}^n \), then for any real numbers \( \alpha \) and \( \beta \), \( \alpha K_1 + \beta K_2 \) is also convex. More generally, \( TK \) is convex if \( T \) is an arbitrary linear mapping of \( \mathbb{R}^n \) onto itself.

**Definition:** Support function of \( K \)

Let \( K \) be a convex set in \( \mathbb{R}^n \) and let \( K \) contain the origin \( 0 \).

For each \( x \in \mathbb{R}^n \), let \( I(x) = \{ a : a > 0, x/a \in K \} \) and \( k(x) = \inf a : a \in I(x) \) then \( k(x) \) is called the support function of \( K \).

**Examples:** Consider the real line \( \mathbb{R}^1 \); then any convex set in \( \mathbb{R} \) is an interval

\[
\begin{array}{cccccccccccc}
-7 & -6 & -5 & -4 & -3 & -2 & -1 & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\
\end{array}
\]

\( A = [-1, 2] \), (where \([...]\) denotes a closed interval)

A support function for \( A \) may be defined as follows:

\( \forall x \in \mathbb{R}^n \), let \( I(x) = [a : |x/a| < \text{dist}(0,2)] \);

\( \inf I(x) = k(x) \) take \( x = 8 \), then \( k(8) = 8/2 = 4 \).

**Lemma 2:** Let \( K \) contain 0 and \( k \) be its support function. Then:
(1) \( k(x) \geq 0 \)

(2) \( k(x) < \infty \)

(3) \( k(ax) = ak(x) \) for \( a > 0 \)

(4) if \( x \in K \) then \( k(x) \leq 1 \)

(5) \( k(x + y) < k(x) + k(y) \)

(6) The set of internal points of \( K \) is characterized by the condition \( k(x) < 1 \); the set of boundary points by the condition \( k(x) = 1 \); and the set of points outside of \( K \) by the condition \( k(x) > 1 \).

(7) For any linear transformation \( T \) from \( \mathbb{R}^n \) into \( \mathbb{R}^n \), \( k(Tx) = k(x) \).

**Hahn-Banach Theorem:** Let the real function \( k \) on the real linear space \( \mathbb{R}^n \) satisfy the conditions:

\[
 k(x + y) \leq k(x) + k(y) , \quad k(\alpha x) = \alpha k(x) ; \quad \alpha \geq 0 , \ x, y \in \mathbb{R}^n .
\]

Let \( f \) be a real linear function on a subspace \( H \) of \( \mathbb{R}^n \) with

\[
 f(x) \leq k(x) , \ x \in H .
\]

Then there is a real linear function \( F \) on \( \mathbb{R}^n \) for which

\[
 F(x) = f(x) , \ x \in H , \text{ and } F(x) < k(x) , \ x \in \mathbb{R}^n .
\]

**Definition:** A linear function \( F \) separates sets \( M \) and \( N \) if and only if there exists a real number \( c \) such that \( F(m) > c \) and \( F(n) < c \) for all \( m \in M, n \in N \).

**Theorem I** The Basic Separation Theorem

Let \( M \) and \( N \) be disjoint convex subsets of \( \mathbb{R}^n \) and let \( M \) have an internal point. Then there exists a non-zero linear function \( F \) which separates \( M \) and \( N \).

**Theorem II** A linear function separates \( M \) and \( N \) in \( \mathbb{R}^n \) if and only if it separates the subsets \( M-N \) and zero in \( \mathbb{R}^n \). It is a separation function of \( M \) and \( N \).
Proof of Theorem I

We may let $M$ contain the origin 0 of $\mathbb{R}^n$ without loss of generality. Assume $M$ contains the origin.

Let $p$ be any point of $N$ then $-p \in (M - N)$, $0 \notin (M - N + p)$, but $p \notin (M - N + p)$ because $M - N$ does not contain 0, since $N \cap M = \emptyset$.

Let $K = M - N + p$, and $k$ be a support function of $K$. Then $k(p) > 1$ and $k(x) < 1$ for $x \notin K$.

If we put $F_1(\alpha p) = \alpha k(p)$, then $F_1$ is a linear function defined on the one-dimensional subspace of $\mathbb{R}^n$ which consists of real multiples of $p$.

Moreover $F_1(\alpha p) \leq k(\alpha p)$ for all real $\alpha$. Since $\alpha \geq 0$ we have $F_1(\alpha p) = k(\alpha p)$, while for $\alpha < 0$ we have $F_1(\alpha p) = \alpha F_1(p) < 0 \leq k(\alpha p)$.

Using the Hahn-Banach Theorem, $F_1$ can be extended to a linear function $F$ on $\mathbb{R}^n$ such that $F(x) \leq k(x)$ for all $x \in \mathbb{R}^n$. It follows that $F(x) \leq k(x) < 1$ for $x \in K$ while $F(p) > 1$. Thus $F$ separates $K$ and $\{p\}$. Therefore $F$ separates $N - N$ and $\{0\}$ and hence $F$ separates $M$ and $N$.

2. CLASSIFYING A PATTERN INTO ONE OF TWO CLASSES

Let $M$ and $N$ be two sets of real valued attribute vectors in $n$-dimensional Euclidian space:

$$M = \{m = (m_{11}, m_{12}, \ldots, m_{1n}) : m_{1i} \in \mathbb{R}^n \text{ for } i = 1, \ldots, l\}$$

$$N = \{n = (n_{11}, n_{12}, \ldots, n_{1n}) : n_{1i} \in \mathbb{R}^n \text{ for } i = 1, \ldots, k\}$$

and denote their convex hulls by $Co(M)$ and $Co(N)$, respectively.

Suppose $Co(M)$ has an internal point, and $Co(M) \cap Co(N) = \emptyset$; that is the convex hulls of $M$ and $N$ are disjoint, and also assume they are bounded.

We may assume a priori that $M$ contains the origin 0 of $\mathbb{R}^n$, and that $N$ intersects the $e_1$-axis. For, letting $a$ be any point in $M$, the set...
\[ M-a = \{ x : x = m_i - a, \forall m_i \in M \} \] contains the origin 0.

Next, by a suitable rotation of the coordinate system, \( N-a \) may be made to intersect with the \( e_1 \)-axis. We may assume \( M \) has the origin and \( N \) intersects with the \( e_1 \)-axis. If the original \( M \) and \( N \) are disjoint convex sets, then, by Lemma 1, the new \( M \) and \( N \) are disjoint convex sets also.

Hereafter, we simplify the notation by writing \( M \) for \( \text{Co}(M) \), \( N \) for \( \text{Co}(N) \).

The problem can be dealt with in two steps:

1. What conditions on \( M \) and \( N \) insure that there exists a linear separation function in \( \mathbb{R}^n \) which separates \( M \) and \( N \)?

2. If \( M \) and \( N \) are separated by the separation function \( f \), how can we find \( f \)?

The answer to question [1] is given by the Hahn-Banach Theorem, and may be summarized as follows: \( f \) separates \( M \) and \( N \) if and only if \( t \) separates \( M-N + p \) and \( \{ p \} \) for \( \forall p \in N \). First define a support function of \( K \) and then construct the linear separation function on \( \mathbb{R}^n \) which satisfied

\[ f(x) \leq k(x) \quad \forall x \in \mathbb{R}^n, \]

where \( k(x) \) is the support function. Since, for any element \( x \) of \( k \),

\[ k(x) < 1, \quad f(m-n) < 1-f(p), \quad \forall m \forall n \quad (m-n) \in (M-N). \]

But \( f \) was chosen so that \( f(p) > 1 \). Therefore, \( f(m-n) < f(m) - f(n) < 0 \) \( f(m) < f(n) \).

Therefore, by the continuity axiom that exists a \( c \) such that \( f(m) < c < f(n) \).

The answer to question [2] (i.e. how \( f \) is to be found) is as follows: One way to solve this problem is to construct \( F_i(x_i) \) on \( \mathbb{R}^i \) inductively for \( i = 1, 2, \ldots, n \).
We assumed that $M$ contains 0 and that $N$ has a non-empty intersection with the $e_1$-axis. So we may take $p \in N \cap [e_1$-axis], and let $K = M - N + p$. Then, as argued earlier, $K$ contains 0 but not $p$. For reasons that will subsequently become apparent, it is convenient to shrink $K$ so that it lies within a sphere with center at 0 and radius $\frac{\sqrt{2}}{2}$. This can be accomplished by applying a transformation

$$S = \begin{bmatrix} \frac{\sqrt{2}}{2L} & 0 \\ 0 & \frac{\sqrt{2}}{2L} \end{bmatrix}$$

where $L = \max (|x|)$ on $\mathbb{R}^n$.

I. Construction of $R^1$ and $F_1(x_1)$.

Define $R^1 = \{\alpha_1 p\}$ where $\alpha_1$ is a real number. Then $R^1$ will be the $e_1$-axis. Define $F_1(x_1) = F_1(\alpha_1 p) = \alpha_1 c_1$. But according to the proof of Theorem 1, $c_1$ may be chosen so that

$$F_1(x_1) \leq k(x_1) \quad \forall x_1 \in R^1$$

$$F_1(x_1) < 1 \quad \text{if } x_1 \in k \cap R^1$$

and $F_1(p) > 1$.

Therefore we choose $c_1$ as follows:

$$1 < c_1 \leq k(p)$$

II. Construction of a separation function $F_2$ on $R^2$ and $R^2$.

In $R^1$, $p$ was a single base of $R^1$ instead of $e_1 = (1, 0, \ldots, 0)$ and obviously $R^1$ does not contain $e_2 = (0, 1, 0, \ldots, 0)$ which is an element of the base of $n$-dimensional vector space. Every vector in the subspace $R^2$ spanned by $R^1$ and $e_2$ has a unique representation in the form $x_1 + \alpha_2 e_2$ with $x_1$ in $R^1$. We want to get a real linear extension $F_2$ of $F_1$ for which the inequality holds $F_2(x_2) \leq k(x_2)$ for $\forall x_2 \in R^2$ and also $F_2(x_1) = F_1(x_1)$. 
for $\forall x_1 \in \mathbb{R}^1$, hence $F_2(x_2) < 1$ if $x_2 \in K \cap \mathbb{R}^2$.

For any constant $c_2$, the function $F_2$ defined on $\mathbb{R}^2$ by the equation $F_2(x_1 + \alpha_2 e_2) = F_1(x_1) + \alpha_2 c_2$ is a proper extension of $F_1$. The desired function $F_2$ we will be able to get by choosing a suitable $c_2$ which satisfies the inequalities given by the Hahn-Banach Theorem:

$$-k(-e_2 - x_1) - F_1(x_1) \leq c_2 \leq k(e_2 + x_1) - F_1(x_1) \quad \forall x_1 \in \mathbb{R}.$$ 

The Hahn-Banach Theorem is of basic importance in the analysis of questions concerning the existence of a continuous linear function. The answer to question [1] was given by this theorem when we want to classify a pattern into one of two classes. But the answer to the question [2] is not so easy: the inequalities given in the Hahn-Banach Theorem do not help us in practice, because the value of $k(e_2 + x_1)$ cannot be determined, due to the possibly irregular shape of the convex set $K$. It was, in fact, for the purpose of replacing these by a pair of inequalities that we applied the transformation $S$ on $\mathbb{R}^n$ so that $K$ would be confined within a circle of radius $\sqrt{2}/2$ centered at 0. The radius was chosen on the basis of the following considerations.

In Figure 1, $TT'$ is the tangent from point $e_2 = (0,1,0,0,\ldots)$ to the confining circle, $T'$ being its $e_1$-axis intercept. The radius $r$ was chosen so that $OT' = 1$. It then follows from Figure 1 that

$$r^2 = 1 - \frac{(TT')^2}{2}$$

$$= 1 - \left(\frac{\sqrt{2}}{2}\right)^2 = \frac{1}{2},$$

and hence $r = \sqrt{2}/2$. 
But $K = y$, $k(e^2 + x) = 1 + (\alpha^p)$.

Since $y < OT' = 1$, it follows that $\frac{1}{y} > 1$, and hence $k(e_2 + x_1) = 1 + \frac{\alpha^p}{y} > 1 + |\alpha^p|$. Since $k(x_1 + e_2) > F_1(x_1)$ according to the construction of $F_1(x_1)$, $F_1(x_1 - y_k)$ is smaller than $k(x_1 - y_k)$ for any $s_1$ and $y_1$ in $\mathbb{R}^1$. From the property of the support function $k$, we get
Therefore, by virtue of the linearity of $F_1$, we get the Hahn-Banach inequality on $\mathbb{R}^2$

$$-k(-y_1 - e_2) - F_1(y_1) \leq k(x_1 + e_2) - F(x_1).$$

This inequality is satisfied for arbitrary choices of $x_1, y_1$ in $\mathbb{R}^1$, and the variable $y_1$ in the left side of this is independent of the choice of $x_1$ in the right side of the inequality. Therefore we can say that there always exists a constant real number $c_2$ between them. That is,

$$-k(-x_1 - e_2) - F_1(x_1) \leq c_2 \leq k(x_1 + e_2) - F(x_1).$$

If $1 + |\alpha_1 p| > F_1(x)$, then there must always exist a constant $c_2$ such that

$$c_2 \leq 1 + |\alpha_1 p| - F_1(x_1) \leq k(x_1 + p_2) - F_1(x_1)$$

and

$$-k(-x_1 - e_2) - F_1(x_1) \leq c_2.$$

For such a constant $c_2$, the linear extended function of $F_1(x_1)$ on $\mathbb{R}_1$ to $\mathbb{R}^2$ will be given by the proper form

$$F_2(x_2) = F_1(x_1) + c_2 x_2,$$

where $x_2 = x_1 + \alpha_2 e_2$.

This new real linear function on $\mathbb{R}^2$ will take a smaller value than the value of the support function $k(x_2)$. That is, $F_2(x_2) \leq k(x_2)$.

Thus, the second inequality in (2) may be replaced by

$$c_2 \leq 1 + |\alpha_1 p| - \alpha_1 c_1 \leq k(p_2 + x_1) - F_1(x_1) \quad (2a)$$

Similarly, the first inequality in (2) is replaced by

$$-k(-x_1 - e_2) - F_1(x_1) \leq -1 - |\alpha_1 p| \leq c_2 \quad (2b)$$

From (2a) and (2b) we get

$$-k(-e_2 - x_1) - F_1(x_1) > -1 - |\alpha_1 p| - \alpha c_1 \leq c_2 \leq 1 + |\alpha_1 p| - F_1(x_1) < k(e_2 + x_1) - F_1(x_1). \quad (2)$$
Thus, if we choose $c_2$ to satisfy the inequalities (2), then for any $x_2$ in $\mathbb{R}^2$, the value of $F_2(x_2)$ is less than the value of $k(x_2)$ by the Hahn-Banach Theorem. That is, $F_2(x_2) < k(x_2)$ for all $x_2 \in \mathbb{R}^2$.

Therefore, if $x_2 \in K \cap \mathbb{R}^2$, then $F_2(x_2) \leq k(x_2) < 1$.

\[ F_2(m - n + p) - F_2(m) - F_2(n) < 1 - F_2(p) \]

\[ F_2(p) = F_1(p) > 1 \]

\[ 1 - F_2(p) < 0 \]

\[ F_2(m) > F_2(n) \]

That is, $F_2$ separates $M \cap \mathbb{R}^2$ and $N \cap \mathbb{R}^2$.

Similarly, we can construct $\mathbb{R}^3$ including $\mathbb{R}^2$ as its subspace and define $F_3$ on $\mathbb{R}^3$ which is linear, takes the same value as $F_2$ on $\mathbb{R}^2$, and separates $M \cap \mathbb{R}^3$ and $N \cap \mathbb{R}^3$. That is,

\[ F_3(x_3) = F_3(x_2 + \alpha_3 e_3) = F_2(x_2) + \alpha_3 c_3 \]

where $x_2 \in \mathbb{R}^2$

\[ x_3 = x_2 + \alpha_2 e_3 = x_1 + \alpha_2 e_2 + \alpha_3 e_3 \]

and $e_3 = (001...)$.

\[ F_3(x_3) \leq k(x_3) \text{ for all } x_3 \in \mathbb{R}^3 \]

where $c_3$ was chosen from the region between

\[ -1 - |\alpha_1 p + \alpha_2 e_2| - F_2(x_2) \quad \text{and} \quad +1 + |\alpha_1 p + \alpha_2 e_2| - F_2(x_2) \]

or

\[ c_3 = c_3 - F_2(x_2) \quad \text{where} \quad -1 - |\alpha_1 p + \alpha_2 e_2| \leq 1 + |\alpha_1 e + \alpha_2 e_2| \]

We can construct $\mathbb{R}^n$ and $F_n$ successively,

\[ F_n(x_n) = F_n(x_{n-1} + \alpha_n e_n) = F_{n-1}(x_{n-1}) + \alpha_n c_n \]

\[ = F_{n-2}(x_{n-2}) + \alpha_{n-1} c_{n-1} + \alpha_n c_n \]

\[ = \alpha_1 c_1 + \alpha_2 c_2 + \ldots + \alpha_n c_n \]
where

\[-k(-x_{i-1} - e_i) - F_{i-1}(x_{i-1}) \leq C_i \leq k(x_{i-1} + e_i) - F_{i-1}(x_{i-1})\]

and

\[x_n = x_{n-1} + \alpha_n e_n = x_{n-2} + \alpha_{n-1} e_{n-1} + \alpha_{n-2} e_{n-2} + \ldots = \alpha_1 p + \ldots + \alpha_n e_n\]

where

\[e_i = (0, \ldots, 0, 1, 0, \ldots, 0), \alpha_i \in \mathbb{R}.

It is clear that \( \mathbb{R}^n = \mathbb{R}^n \), thus \( F_n \) can be considered the required separation function of \( M \) and \( N \) or \( \mathbb{R}^n \). Therefore for any element \( x \) of \( \mathbb{R}^n \),

\[F_n(x) = F_{n-1}(x_{n-1}) + \alpha_n C_n\]

if \( \alpha_n = 0 \),

\[F_n(x) = F_{n-1}(x_{n-1}) \leq k(x_{n-1})\]

if \( \alpha_n > 0 \), replace \( x_{n-1} \) by \( x_{n-1}/\alpha_n \),

and we have

\[F_{n-1}(x_{n-1}) + \alpha_n C_n < F_{n-1}(x_{n-1}) + \alpha_n (kC_n + \frac{x_{n-1}}{\alpha_n}) - F_{n-1}(\frac{x_{n-1}}{\alpha_n})\]

\[= \alpha_n k\left(e_n + \frac{x_{n-1}}{\alpha_n}\right) = k(\alpha_n + x_{n-1})\]

\[\therefore F_n(x) < k(x_n)\]

\( F_n \) is the required linear separation function, because for \( \forall x_n = m - n + p \),

\[F_n(x) = F_n(m - n + p) = F_n(m) - F_n(n) + F_n(p) \leq k(x_n) < 1\]

\[F_n(m) - F_n(n) \leq 1 - F_n(p)\]

Since \( F_n(p) = F_1(p) = C_1 > 1, 1 - F_n(p) < 0\)

\[\therefore F_n(m) < F_n(n)\] for any \( m \in M \) and \( n \in N \).

Therefore, \( \max F_n(m) < \min F_n(n) \)

\[m \in M \quad n \in N\]

Therefore, there exists a \( C \) such that \( \max F_n(m) < C < \min F_n(n) \)

\[m \in M \quad n \in N\]
Let us say we want to know whether a given vector $x$ belongs to $M$ or to $N$. We need only calculate the value of $F_n(x)$ and compare with $C$; i.e.,

- if $F_n(x) < C$ then $x \in M$
- if $F_n(x) > C$ then $x \in N$

The steps are summarized in Table 1.

<table>
<thead>
<tr>
<th>Table 1. Summary of Method: I. Two Classes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. M, N Given in $\mathbb{R}^n$, $0 \in M$, $M \cap N = \emptyset$, convex, $N \cap {x: y = 0} \neq \emptyset$.</td>
</tr>
<tr>
<td>2. $K = M - N + P \forall P \in N \cap {x\text{-axis}}$ then $K \ni 0$, $P \notin K$.</td>
</tr>
<tr>
<td>3. Construct $R^1 = {\alpha_1 p}; \alpha_1$ is any real number.</td>
</tr>
<tr>
<td>4. Define $F_1$ on $R^1$, $F_1(x_1) = F_1(\alpha_1 p) = \alpha_1 c_1$ where $1 &lt; c_1 \leq k(p)$</td>
</tr>
<tr>
<td>$k(p)$ is a support function of $K \ni k(K) \leq 1$, $k(\overline{K}) &gt; 1$.</td>
</tr>
<tr>
<td>5. Define $F_2$ on $R^2$, $R^2 = {\alpha_1 p + \alpha_2 e_2} = {x_1 + \alpha_2 e_2}$ $e_2 = (0, 1, \ldots)$</td>
</tr>
<tr>
<td>$F_2(x_2) = F_2(x_1 + \alpha_2 e_2) = F_1(x_1) + \alpha_2 c_2$</td>
</tr>
<tr>
<td>$-1 -</td>
</tr>
<tr>
<td>then $F_2(x_1) = F_1(x_1)$ $F_2(x_2) \leq k(x_2)$ $\forall x_2 \in R^2$.</td>
</tr>
<tr>
<td>6. Define $F_3$ on $R^3$, ..., $F_n$ on $R^n$</td>
</tr>
<tr>
<td>$-1 -</td>
</tr>
<tr>
<td>$1 +</td>
</tr>
<tr>
<td>7. Then $F_n$ separates $M$ and $N$, $F_n(M) &lt; F_n(N)$.</td>
</tr>
<tr>
<td>8. Max $F_n(M) = S_1$, min $F_n(N) = S_2$ then for any unknown vector</td>
</tr>
<tr>
<td>if $F_n(x) &lt; \frac{S_1 + S_2}{2}$ then $x \in M$.</td>
</tr>
<tr>
<td>if $F_n(x) &gt; \frac{S_2 + S_1}{2}$ then $x \in N$.</td>
</tr>
</tbody>
</table>
Let $M$ be a square region whose sides are of unit length.

Let $N$ be the interval $[3, 4]$.

Choose $p$ in $N$ as $p = 3$.

**Figure 3.**

1. $M - N + p = \left\{ z \mid z = m - n + 3 \right\} = \left\{ z \mid z = (m_1, m_2) - (n_1, 0) + (3, 0) \right\}$

$-N + 3 = [-1, 0]$

$K = M - N + 3$ is a rectangular region which is represented by dots (\ldots).

2. Since maximum $(|z|) = \sqrt{\left(\frac{3}{2}\right)^2 + \left(\frac{1}{2}\right)^2} = \sqrt{\frac{10}{2}} > \frac{\sqrt{2}}{2}$,

we have to multiply the base of $R^n$ by

$$\frac{1}{\frac{2}{\sqrt{2}} \cdot \frac{\sqrt{10}}{2}}$$

That is, consider the transformation $T = \begin{bmatrix} \sqrt{5} & 0 \\ 0 & \sqrt{5} \end{bmatrix}$, and $R^n = TR^n$

$TK = \sqrt{5} M - \sqrt{5} N + 3\sqrt{5}$, $e_1 = (1, 0)$, $e_2 = (0, 1)$

**Figure 4.**
3. Determine $F_2$.

Let $p = \frac{3}{\sqrt{3}}$, define $R^1$ and $F_1$ as follows:

$$R^1 = \{\alpha_1 p\} = \{x_1\}$$

$$F_1(x_1) = \alpha_1 C_1$$

where $1 > C_1 \leq k(p) \frac{3}{1/2} = 6$.

If we take $C_1 = 6$, then

$$F_1(x_1) = 6\alpha_1.$$

Take any element $m-n+p$ in $K \cap R^1$. Then $F_1$ has to classify $m$ and $n$ in $K \cap R^1$. In order for $F_1$ to be a separation function of $M \cap R^1$ and $N \cap R^1$, first decide the number $c$ such that

$$\max F_1(m) \leq c \leq \min F_1(n).$$

$m \in M \cap R^1$, $n \in N \cap R^1$

Since

$\max F_1(m) = 1$ when $\alpha_1 = \frac{1}{6}$

$\min F_1(n) = 6$ when $\alpha_1 = 1,$

we can take any $c$ from the interval $(1, 6)$ to construct $F_2$ on $R^2$.

We now proceed to define $F_2$ and $R^2$ as follows:

$$F_2(x_2) = F_1(x_1) + \alpha_2 c_2$$

where $-1 < c_2 \leq 1$

$$c_2 = c'_2 - F_1(x_1) = c'_2 = 6\alpha_1.$$

If we take $c'_2 = 6\alpha_1$, then

$$F_2(x_2) = 6\alpha_1 + \alpha_2 (1-6\alpha_1).$$

To let $F_2$ separate $M$ and $N$, we determine $c$ so that

$$\max F_2(m) \leq c \leq \min F_2(n).$$

$n \in N$
Therefore it is enough to find what \( \max c_2 \alpha_2 \) is, i.e., \( \max \alpha_2(1-6\alpha_1) \).

Since 
\[
\max (\alpha_2) = \frac{1}{2}, \\
(x_1 + \alpha_2 e_2) \in M \\
\max \alpha_2(1-6\alpha_1) = \frac{1}{2} (1 - 1) = 0
\]

Therefore, \( \max F_2(m) = 1 \)
\[
m \in M
\]

Further, since \( \min (\alpha_2) = 0, \)
\[
x_1 + \alpha_2 \in N \\
\min F_2(n) \geq \min F_1(n) + \alpha_2(1 - 6\alpha_1) \\
n \in N
\]

Therefore \( c \) must be taken from the interval \((1, 6)\).

Thus, we finally get the separation function which we want.

We show that this \( F_2 \) is indeed the separation function of \( M \) and \( N \).

For any element \( m \) in \( M \), \( m = \alpha_1 p + \alpha_2 e_2 \)
\[
\begin{align*}
-\frac{1}{2} & \leq \alpha_1 \leq \frac{1}{6} \\
-\frac{1}{2} & \leq \alpha_2 \leq \frac{1}{2}
\end{align*}
\]

\[
F_2(m) = 6\alpha_1 + \alpha_2(1 - 6\alpha_1)
\]

\( F_2(m) \leq 1 < c \).

Next let us check an element of \( N \). For any element \( n \) in \( N \)
\[
n = \alpha_1 p + 0e_2
\]

\( F_2(n) = 6\alpha_1 > 6 > c. \)

3. CLASSIFYING PATTERNS INTO ONE OF A FINITE NUMBER OF CLASSES

In the previous section, we have demonstrated the existence of a real linear function which separates the convex and bounded disjoint sets \( M \) and \( N \) in \( \mathbb{R}^n \), and shown how to find such a separation function. In this section we are going to discuss the existence of a separation function and the extension of the method to obtain this function in the case when there are a
finite number of subsets, (say at most \( n \) in \( \mathbb{R}^n \)) which we want to discriminate from one another.

We shall use the letters \( A_i \), \( i = 1 \ldots n \) instead of using the letters \( M, N \) as the sets to be discriminated in \( \mathbb{R}^n \) hereafter.

We assume these sets \( A_i \), \( i = 1, \ldots n \) must be mutually disjoint, convex and bounded (not necessarily finite sets); that is, \( A_i \cap A_j = \emptyset \) for \( i \neq j \), the diameters of \( A_i \) are finite numbers, where the diameter of \( A_i = \text{dia} (A_i) = \sup (|a_{i} - a'_{i}|) \)

\[ a_i, a' \in A_i \]

**Definition:** A real linear function \( F_n \) defined on \( \mathbb{R}^n \) separates the set \( A_i \) from \( \bigcup_{j \neq i} A_j \) if and only if \( F_n \) separates \( A_i \) and \( A_j \) for all \( j \) such that \( i \neq j \).

**Theorem:** If the sets \( A_1, A_2, \ldots, A_n \) are convex, bounded and mutually disjoint subsets of \( \mathbb{R}^n \), then there exists a non-zero real linear function \( F_n \) on \( \mathbb{R}^n \) which separates \( A_1 \) and \( \bigcup_{i \neq 1} A_i \).

In order to prove this theorem, we consider a dummy circle, set \( C' \), whose radius is a sufficiently small real number \( \epsilon \) and the distances from any one of the given sets \( A_i \), \( i = 1, \ldots, n \) are strictly greater than the total sum of diameters of \( A_i \) and \( \epsilon \); i.e., distance \( (C', A_i) \geq \sum_{i=1}^{n} \text{dia} (A_i) + \epsilon \), \( C' \cap A_i = \emptyset \) for \( i = 1, \ldots, n \).

Without loss of generality, we can assume the center of this circle to be the origin of \( \mathbb{R}^n \). In the previous section the support function of the convex set \( K = M - L + p \) was used to obtain the separation function of \( M \) and \( L \) in \( \mathbb{R}^n \). In this section we consider a circle \( K \) concentric to \( C' \) whose radius is \( r = \sum_{i=1}^{n} \text{dia} (A_i) + \epsilon \) as the convex
set which has the same role as \( M - L + p \) in section 2.

**Property 1:** The set \( K \) and \( A_i \), \( i = 1, 2, \ldots, n \) are disjoint,

\[ K \cap A_i = \emptyset \quad \text{for} \quad i = 1, \ldots, n. \]

**Property 2:** Let \( p_i \) be an arbitrary point in \( A_i \) and let

\[ A_i - p_i = \{ x_i : x_i = y_i - p_i \quad \forall y_i \in A_i \} \]

\[ -A_i = \{ w_i : w_i = -y_i \quad \forall y_i \in A_i \} \]

\[ -A_i + p_i = \{ v_i : v_i = -y_i + p_i \quad \forall y_i \in A_i \}. \] Then \( K \) contains \( A_i - p_i, -A_i + p_i, \quad A_i - p_i + A_i - p_i, \) and \( C' - A_i + p_i \).

**Property 3:** Each \( p_i \) is not in \( A_i \); \( p_i \notin A_i \).

**Property 4:** Let \( k(x) \) be a support function of \( K \).

\( k(x) \) will be given by \( k(x) = |x|/r \) where \( |x| \) is the absolute value of the vector \( x \) in \( \mathbb{R}^n \).

\( k(x) < 1 \) characterizes the internal point of \( K \),

\( k(x) > 1 \) characterizes the outside point of \( K \), and

\( k(x) = 1 \) characterizes the boundary point of \( K \).

This function \( k(x) \) satisfies the all properties of the support function represented in the previous section.

We shall complete the proof by constructing the non-zero real linear function \( F_n \) on \( \mathbb{R}^n \) which separates the sets \( A_i \) and \( \bigcup_{i \neq 1} A_i \).

We first construct a one-dimensional subspace \( \mathbb{R}^1 \) in \( \mathbb{R}^n \) and the linear function \( F_1 \) on \( \mathbb{R}^1 \) which separates \( A_1 \cap \mathbb{R}^1 \) and \( C' \cap \mathbb{R}^1 \).

Suppose \( p_1 \) is an arbitrary point in \( A_1 \). Define \( \mathbb{R}^1 = \{ \alpha_1 p_1 \} = \{ x_1 \} = \{ (\alpha_1) \} \) where \( \alpha_1 \) is a real number, then \( \mathbb{R}^1 \) is the straight line which goes through the origin in \( \mathbb{R}^n \) and the point \( p_1 \). \( p_1 \) constitutes a
single-element of base of the subspace $\mathbb{R}^1$ in $\mathbb{R}^n$, and $(\alpha)$ is a component of a vector $x_1$. Next, as we did in section 2, we define a linear function $F_1$ on $\mathbb{R}^1$ such that $F_1(x_1) = F_1(\alpha_1p_1) = \alpha_1c_1$ where $c_1$ is a real constant. If we choose $c_1$ as $F_1(p_1) = c_1 = k(p_1)$, then the newly defined linear function $F_1(x_1)$ on $\mathbb{R}^1$ will have a smaller value on the domain $\mathbb{R}^1$ than the value of $k(x)$ on $\mathbb{R}^1$, because $F_1(x_1) = \alpha_1c_1 = \alpha_1k(p_1) - k(\alpha p_1) = k(\alpha_1)$ for $\alpha_1 \geq 0$ while for $\alpha_1 < 0$ while for $\alpha_1 \geq 0$ we have $F_1(\alpha_1p_1) = \alpha_1F_1(1) < 0 < k(\alpha_1p_1)$; i.e., $F_1(x_1) < k(x)$ and $F_1(p_1) > 1$. Moreover if $x_1 \in K \cap \mathbb{R}^1$ then $F_1(x_1) < k(x_1) < 1$.

The value of $F_1(x_1)$ is always positive because, for any element $c_0 - a_1 + p_1$ in the set $C - A_1 + p_1 \subseteq K$ in $\mathbb{R}^1$ where $c_0$ is taken from the dummy set circle $C'$ and $a$ from $A_1$, we have $F_1(c_0 - a_1 + p_1) < 1$. By the linearity of $F_1$ we have $F_1(c_0) - F_1(a_1) < 1 - F_1(p_1) < 0$. Therefore $F_1(c_0) < F_1(a_1)$. Since $F_1$ is a linear continuous function and the set $A_1$ is convex, we get $0 < F_1(a_1)$ for $\forall a_1 \in A_1$.

The next step is to construct a subspace $\mathbb{R}^2$ of $\mathbb{R}^n$ which includes $\mathbb{R}^1$ as its subspace, then define the proper extended function $F_2$ of $F_1$ on $\mathbb{R}^2$ which will separate the sets $A_1$ and $A_2$ in $\mathbb{R}^2$ and has the same value as the value of $F_1$ on $\mathbb{R}^1$, always $F_2(x_2) < k(x_2)$ where $x_2 \in \mathbb{R}^2$.

First, we choose an arbitrary point $p_2$ in $A_2$, then since the sets $A_1$ and $A_2$ are disjoint, the point $p_1$ in $A_1$ and the point $p_2$ in $A_2$ are different, $p_1 \neq p_2$, so if we choose $p_2$ as the other member of the base to span $\mathbb{R}^2$ together with $p_1$, then any element in $\mathbb{R}^2$ will be expressible in the form $x_2 = \alpha_1p_1 + \alpha_2p_2$ or $x_2 = (\alpha_1, \alpha_2)$ where $\alpha_1$ and $\alpha_2$ are real numbers. We must note that this newly extended two-dimensional space
$\mathbb{R}^2$ of $\mathbb{R}^1$ in $\mathbb{R}^n$ might not have an orthogonal coordinate system although the original $\mathbb{R}^n$ has.

If we put $F_2(x_2) = F_1(x_1) + \alpha_2c_2$, then $F_2(\alpha_1p_1 + \alpha_2p_2) = F_2(\alpha_1p_1) + F_2(\alpha_2p_2) = \alpha_1F_2(p_1) + \alpha_2F_2(p_2)$ by linearity of $F_2$.

Therefore we get $F_2(\alpha_1p_1) = F_1(x_1)$ and $F_2(p_2) = c_2$. Since we have not assigned any restriction on $c_2$ except that it must be a real number, we may choose $c_2$ so that $F_2(x_2)$ will be the desired separation function of $A_1$, $A_2$, and $C'$ in subspace $\mathbb{R}^2$. That is, by the Hahn-Banach Theorem if we choose $c_2$ as a real number which satisfied the inequalities

$$-k(-x_1 - p_2) - F_1(x_1) \leq c_2$$

and

$$c_2 \leq k(x_1 + p_2) - F_1(x_1)$$

for any $x_1$ in $\mathbb{R}^1$, then the value of the function $F_2(x_2)$ will be always smaller than the value of the given support function $k(x_2)$ of $K$ on $\mathbb{R}^2$. Moreover, if $F_2(x_2) \leq k(x_2)$ for any $x_2$ in $\mathbb{R}^2$, then

$$F_2(a_2 - p_2 - a_1 + p_1) \leq k(a_2 - p_2 - a_1 - p_1) < 1$$

where $(a_2 - p_2 - a_1 + p_1) \in (A_2 - p_2 - A_1 + p_1) \cap \mathbb{R}^2 \subset K \cap \mathbb{R}^2$.

and

$$F_2(a_2) - F_2(a_1) < 1 - F_2(p_1) + F_2(p_2)$$

and

$$F_2(a_2) - F_2(a_1) < 1 + c_2 - c_1.$$

In order to let $F_2$ separate the sets $A_1$ and $A_2$, we have to choose $c_2$ to be smaller than $c_1 - 1$, that is, $c_2 < c_1 - 1$.

Therefore we get

$$F_2(a_2) < F_2(a_1)$$

for any $a_2$ in $A_2$, $a_1$ in $A_1$ on the subspace $\mathbb{R}^2$. 
The one thing we have to check is whether or not there exists such a $c_2$ which satisfies both the Hahn-Banach inequalities and $c_2 < c_1 - 1$. Since $c_1$ is greater than 1, there always exists $c_2$ such that

$$k(y_1 + p_2) - F_1(y_1) < c_2 < \min \{k(x_1 + p_2) - F_1(x_1), c_1 - 1\}$$

for all $y_1, x_1 \in \mathbb{R}^1$.

Therefore we can conclude that $F_2(x_2)$ separates the sets $A_1$ and $A_2$ in $\mathbb{R}^2$.

Next, we proceed in the same way as we did before to construct the subspace $\mathbb{R}^2$ and the separation function of $A_1$ and $A_2$ on it until we finally get the space $\mathbb{R}^n$ and the separation function $F_n(x_n)$ of the sets $A_1$ and $A_i$ ($i = 1, \ldots, n$) on it.

$$\mathbb{R}^n = \{x_{n-1} + \alpha_n p_n\} = \{\alpha_1 p_1 + \alpha_2 p_2 + \ldots + \alpha_n p_n\}$$

$$= \{(\alpha_1, \alpha_2, \ldots, \alpha_n)\} = \{x_n\} \text{ where } p_i \in A_i$$

$$F_n(x_n) = F_{n-1}(x_{n-1}) + \alpha_n c_i$$

$$= \alpha_1 c_1 + \alpha_2 c_2 + \ldots + \alpha_n c_n$$

and each $c_i$ has been chosen so that

$$F_i(x_i) \leq k(x_i) \text{ for any } x_i \text{ in } \mathbb{R}^i \text{ (i=1, 2, \ldots, n)}$$

and

$$F_i \text{ separates the sets } A_1 \text{ and } A_i; \quad F(a_i) < F(a_1)$$

where $\forall a_1 \in A_1, \forall a_i \in A_i$.

That is,

$$-k(-x_{i-1} - p_{i}) - F_{i-1}(x_{i-1}) < c_i < \min \{k(x_{i-1} + p_{i}) - F_{i-1}(x_{i-1}), c_{i-1}\}$$

Now we have to show that the domain of $F_n(x_n)$ is equal to the originally given Euclid space $\mathbb{R}^n$. That is, $\mathbb{R}^n = \mathbb{R}^n$. This is clear.
because \( p_i \neq p_j \) for \( i \neq j \) and the number of elements in the new base is \( n \).

**Proof that** \( F_n(x_n) \) **separates** \( A_1 \) and \( A_i \):

Take \( a_i \) in \( A_i \) and \( a_1 \) in \( A_1 \), then

\[
F_n(a_i - p_i - a_1 + p_1) < k(a_i - p_i - a_1 + p_1)
\]

\[
k(a_i - p_i - a_1 + p_1) < 1 \quad \text{since} \quad a_i - a_1 \in A_i - p_i - A_1 \in p_1 \subseteq K
\]

\[
F_n(a_i) - F_n(a_1) < 1 - F_n(p_1) + F_n(p_i)
\]

\[
= 1 - c_i + c_1 < 0
\]

Therefore we get

\[
F_n(a_i) < F_n(a_1) \quad \text{for} \quad \forall a_i \in A_i, \forall a_1 \in A_1
\]

Q. E. D.

This theorem just gives us the separation function \( F_n \) which separates \( A_1 \) and other \( A_i \)'s. But we can construct the separation function \( G_n \) which separates \( A_2, A_3, A_4, \ldots, A_n \) in \( \mathbb{R}^n \) in the same way as the above except that the last point to be added as the base of \( \mathbb{R}^n \) is \( e_n \), which is the \( n \)th element of the coordinate system in \( \mathbb{R}^n \). By continuing the constructions of \( F_n, G_n, \ldots, H_n \), we will be able to separate the \( A_i \)'s one from another.

4. **APPLICATION TO SORTING PATTERNS OF RESPONSES ON THE PLATO INSTRUCTIONAL SYSTEM**

In this section we discuss some applications of our method to the problem of retrieving information from a "dope" tape on which student records are stored. While particular cases can be handled simply by ordinary methods, it would be extremely useful if a general method
were available. A multiple classification procedure such as discussed in the preceding sections provides one approach to a general retrieval system.

To take a very simple case, suppose that students work on a test which comprises $N$ problems, and that a record of each is stored on a magnetic tape. Suppose the author of this test wants to know how many and which students get the "O.K." signal for the $i$th problem within time interval $\Delta t$ and also fails the $j$th problem. If a student gets "O.K." within the $\Delta t$, then the transducer assigns 1, otherwise 0.

We could, of course, comply with the author's request simply by sorting the tape; but we wish to illustrate the method described in this paper. It enables us to find a linear function which separates all $2^N$ possible classes and determines the boundary points between classes. Here we only discuss the simple case when there are three problems in the test. Then the possible combinations of sets to be classified are $2^3 = 8$ in number, and a transducer program could assign to them the points $(0, 0, 0), (0, 0, 1), (0, 1, 0), (1, 0, 0), (0, 1, 1), (1, 0, 1), (1, 1, 0), (1, 1, 1)$ respectively, as follows: All of the items were failed, No. 1 and No. 2 only were failed, and so on. Let us name these singleton sets, $A_0, A_1, A_2, A_3, \ldots, A_7$. We choose $K$ as a circle around the origin $A_0$ whose radius is $\frac{1}{2}$, then $K$ is disjoint for each $A_i$, $i = 1, \ldots, 7$. Let us name $p_1 = (0, 0, 1), p_2 = (0, 1, 0), p_3 = (1, 0, 0)$. (Although in the present, trivial example these are each unique elements of the corresponding sets, in
general, each \( p_i \) may be taken as an arbitrary element of the corresponding \( A_i \).

Since the \( A_i \), for \( i \geq 4 \), are the linear combinations of \( A_1, A_2, A_3 \), with respect to the definition of addition we used in sections 2 and 3, it is enough in this case to consider these three sets in order to classify all eight sets.

In this case, and a good many that arise in this application, classification requires constructing only a single linear function \( F_3 \) on \( \mathbb{R}^3 \). Since \( A_i - p_i \) is 0 for any \( i \), \( F_3(a_i) - F_3(a_j) \) is always smaller than \( F_3(p_j) \). If we choose \( c_i \) and \( c_j \) such that \( F_3(p_i) = c_i < c_j = F_3(c_j) \) then \( F_3 \) will be a single linear function that will classify these given sets.

\[ R^1 \text{ will be defined } R^1 = \{ \alpha_1 p_1 \} = \{ (\alpha_1) \} = \{ x_1 \} \text{ and } \]

\[ F_1(x_1) = k(p_1) = \alpha_1 c_1 = 2\alpha_1 \text{ where } k(p_1) = 1/2 = 2. \]

\[ R^2 \text{ will be given by } R^2 = \{ x_1 + \alpha_2 p_2 \} = \{ (\alpha_1, \alpha_2) \} = \{ x_2 \} \text{ and } \]

\[ F_2(x_2) = F_1(x_1) + \alpha_2 c_2 = 2\alpha_1 + 1/100 \alpha_2. \]

\[ R^3 \text{ will be given by } R^3 = \{ x_2 + \alpha_3 p_3 \} = \{ (\alpha_1, \alpha_2, \alpha_3) \} = \{ x_3 \} \text{ and } \]

\[ F_3(x_3) = F_2(x_2) + \alpha_3 c_3 = 2\alpha_1 + 1/100 \alpha_2 + 2/100 \alpha_3. \]

We selected \( c_2 \) and \( c_3 \) as \( 1/100 \), \( 2/100 \), but there are many possible values for them which satisfy the inequalities we gave in the previous section. So the particular linear function we get here \( F_3(x_3) \) is just one of the family given by the linear function derived previously to separate one set from all others. Although this is a very trivial case, we can apply the same classification method also to the case when the scores are continuously scaled.
Incidentally, if we consider the elementary function of these sets which corresponds with the binary representation, \((0, 0, 0)\) to 0, \((0, 0, 1)\) to 1, \((0, 1, 0)\) to 2, etc., then this function will be one of the many choices consistent with our linear function which separates these sets.

As another example, suppose that we want to retrieve information concerning how many and which students pressed keys in some specified sequential patterns in which the author is interested. More precisely, let \(c_i\) be constant keys, \(V_i\) be sets of variable keys which are taken from a specified domain including the null element 0 meaning no key was pressed. For example, one of the patterns will be denoted by \(A_i = \{c_1, \ldots, c_k, V_1, c_{k+1}, V_2, \ldots, c_n\}\). If we consider the possible sequences in \(A_i\) as elements of the attribute space whose base is the sequential order of key presses, then the first coordinate is the first key press, the second is the second one, so on. Then in this space, each \(A_i\) will be a rectangular parallelepiped of \(n\) dimensions around \(\mathbf{p}_i = (c_1, \ldots, c_k, 0, \ldots, 0, c_{k+1}, 0, \ldots, 0, \ldots, c_n)\). In this way we can apply our method of constructing a linear function to the problem of separating patterns of sequences of student responses. It is expected that we can also apply our method to many cases which would be all but impossible to discriminate only by sorting the data.
REFERENCES


Distribution list as of May 1, 1966 (cont'd.)

1. Commanding Officer
   Office of Naval Research
   1400 Summer Street
   Boston, Massachusetts 02110

2. Director, Naval Research Laboratory
   Technical Information Officer
   Washington, D. C.
   Attn: Code 2000

1. Commander
   Naval Air Development and Material Center
   Johnstown, Pennsylvania 15901

2. Librarian
   U. S. Naval Electronics Laboratory
   San Diego, California 92152

1. Commanding Officer and Director
   U. S. Naval Underwater Sound Laboratory
   Port Trumbull
   New London, Connecticut 06360

1. Librarian
   U. S. Navy Postgraduate School
   Monterey, California

1. Director
   U. S. Naval Air Missile Test Center
   Point Mugu, California

1. Director
   U. S. Naval Observatory
   Washington, D. C.

2. Chief of Naval Operations
   OPNAV
   Washington, D. C.

1. Director, U. S. Naval Security Group
   Attn: G3
   3801 Nebraska Avenue
   Washington, D. C.

2. Commanding Officer
   Naval Ordnance Laboratory
   White Oak, Maryland

1. Commanding Officer
   Naval Ordnance Laboratory
   Corona, California

1. Commanding Officer
   Naval Ordnance Test Station
   China Lake, California

1. Commanding Officer
   Naval Avionics Facility
   Indianapolis, Indiana

1. Commanding Officer
   Naval Training Device Center
   Orlando, Florida

1. U. S. Naval Weapons Laboratory
   Dahlgren, Virginia

1. Weapons Systems Test Division
   Naval Air Test Center
   Patuxent River, Maryland
   Attn: Library

1. Mr. Charles F. Yost
   Special Assistant to the Director of Research
   National Aeronautics and Space Administration
   Washington, D. C. 20546

1. Dr. B. Harrison, Code BEE
   Chief, Electrophysics Branch
   National Aeronautics and Space Administration
   Washington, D. C. 20546

1. Goddard Space Flight Center
   National Aeronautics and Space Administration
   Attn: Library, Documents Section Code 252
   Greenbelt, Maryland 20771

1. NASA Lewis Research Center
   Attn: Library
   21000 Brookpark Road
   Cleveland, Ohio 44135

1. National Science Foundation
   Attn: Dr. John B. Lehmann
   Division of Engineering
   1800 G Street, N. W.
   Washington, D. C. 20530

1. U. S. Atomic Energy Commission
   Division of Technical Information Extension
   P. O. Box 62
   Oak Ridge, Tennessee 37831

1. Los Alamos Scientific Laboratory
   Attn: Research Library
   P. O. Box 1663
   Los Alamos, New Mexico 87544

1. NASA Scientific & Technical Information Facility
   Attn: Acquisitions Branch (TIA/STIF)
   P. O. Box 33
   College Park, Maryland 20740

1. Director
   Research Laboratory of Electronics
   Massachusetts Institute of Technology
   Cambridge, Massachusetts 02139

1. Polytechnic Institute of Brooklyn
   55 Johnson Street
   Brooklyn, New York 11201
   Attn: Dr. Jerome Fouquet
   Research Coordinator

1. Director
   Columbia Radiation Laboratory
   Columbia University
   538 West 120th Street
   New York, New York 10027

1. Director
   Coordinated Science Laboratory
   University of Illinois
   Urbana, Illinois 61801

1. Director
   Stanford Electronics Laboratories
   Stanford University
   Stanford, California

1. Director
   Electronics Research Laboratory
   University of California
   Berkeley, California

1. Director
   Electronic Sciences Laboratory
   University of Southern California
   Los Angeles, California 90007

1. Professor A. A. Dougall, Director
   Laboratories for Electronic and Related Sciences Research
   University of Texas
   Austin, Texas 78712

1. Division of Engineering and Applied Physics
   210 Pierce Hall
   Harvard University
   Cambridge, Massachusetts 02138

1. Aerospace Corporation
   P. O. Box 95805
   Los Angeles, California 90063
   Attn: Library Acquisition Group

1. Professor Nicholas George
   California Institute of Technology
   Pasadena, California

1. Aeronautics Library
   Graduate Aeronautical Laboratories
   California Institute of Technology
   1201 E. California Boulevard
   Pasadena, California 91109

1. Director, TRW Project RAND
   Via: Air Force Liaison Office
   The RAND Corporation
   1700 Main Street
   Santa Monica, California 90406
   Attn: Library

1. The John Hopkins University
   Applied Physics Laboratory
   8641 Georgia Avenue
   Silver Spring, Maryland
   Attn: Boris M. Rovshenoff
   Document Librarian

1. Hunt Library
   Carnegie Institute of Technology
   Schenley Park
   Pittsburgh, Pennsylvania 15213

1. Dr. Leo Young
   Stanford Research Institute
   Menlo Park, California

1. Mr. Henry L. Bachmann
   Assistant Chief Engineer
   Wheeler Laboratories
   122 Chittenden Road
   Great Neck, New York

1. University of Liege
   Electronic Department
   Mathematics Institute
   15, Avenue Des Tilleuls
   Val-Sainte, Liege
   Belgium

1. School of Engineering Sciences
   Arizona State University
   Tempe, Arizona

1. University of California at Los Angeles
   Department of Engineering
   Los Angeles, California

1. California Institute of Technology
   Pasadena, California
   Attn: Documents Library

1. University of California
   Santa Barbara, California
   Attn: Library

1. Carnegie Institute of Technology
   Electrical Engineering Department
   Pittsburgh, Pennsylvania

1. University of Michigan
   Electrical Engineering Department
   Ann Arbor, Michigan

1. New York University
   College of Engineering
   New York, New York

1. Syracuse University
   Department of Electrical Engineering
   Syracuse, New York

1. Yale University
   Engineering Department
   New Haven, Connecticut

1. Airborne Instruments Laboratory
   Newport, New York

1. Bendix Pacific Division
   11600 Sherman Way
   North Hollywood, California

1. General Electric Company
   Research Laboratories
   Schenectady, New York

1. Lockheed Aircraft Corporation
   P. O. Box 504
   Sunnyvale, California

1. Raytheon Company
   Bedford, Massachusetts
   Attn: Librarian
In this paper we develop a method for classifying any object which can be represented as a point in n-space into one of m given subsets by constructing a linear function which is derived from the support function in linear topological space, an approach hitherto unexplored in this connection.

Greenberg and Konheim (1964) discussed this problem of classification scheme by presenting two possible procedures: one is to use pairwise class separation, and the other is to construct m linear functions which separate one of the sets from all the others. The former requires the construction of m(-1)/2 linear functions, and the latter places more stringent restrictions on the m subsets than the former. We show, in this report, how to construct m instead of m(m-1)/2, linear functions under the more liberal assumptions of the first procedure, thus combining the merits of both procedures.

The application of our procedure to a problem in the classification of sequences of student responses on the PLATO teaching system is discussed. In that connection, a possible further development, an even simpler method requiring only one linear function, was explored. The validity of this simpler method could, however, be established only under rather restrictive
Linear Separability  
Pattern Classification  
Support Function  
Separation Function  
Computer-based Instruction