

# 直交空間での教師なし類別について†

瀬 良 豊 士††

## Abstract

It is known that Karhunen-Loève orthogonal system can be applied to unsupervised clustering. And some results are obtained with the method for unsupervised clustering by the K-L system. But, in the case of the practical patterns, the results are not so much given. So, first the method for unsupervised clustering by the K-L system are explained from the basic points of view. Secondly, an algorithm for the unsupervised clustering are given under the method described above. Finally, the reformed patterns, which were proposed by the author, are used as the practical patterns and computer-simulated experiments are carried into effect in order to test the usefulness of the present method.

## 1. 緒言

パタン認識の問題は、対象とするものについての予備知識があるかないかによって次の2つに分けられる。その1つはまえもって知られているクラス集合のいずれかに与えられた未知パターンが含まれるものと決定するパタン認識の問題であり、他の1つは、多数の代表パタン集合から共通なもの同志を同一クラスとしてまとめクラスの集合を形成する認知の問題、すなわちクラスタリングの問題である。ここでは未知パターンベクトルを直交成分に展開する Karhunen-Lóeve 直交系を用いて、分布の未知なパタン集合からクラスの集合を形成するクラスタリングの一方法を考察する。

K-L 系の基本的概念は、1963年飯島<sup>1)</sup>によって自己相関の概念を用いた特徴抽出論で明らかにされ、1965年に Watanabe<sup>2)</sup> が独立な立場から同じ概念を明らかにし、これを Karhunen-Lóeve 直交系と名づけた。1970年 Fukunaga<sup>3)</sup> らは、この K-L 系を用いて、分散が互いに等しい場合の2分類問題を考察し、パタンの入れかえによるアルゴリズムを明らかにし、教師なしの類別を提案した。さらに、1971年富田<sup>5)</sup> らは上記の理論をさらに一般化した教師なし2類別問題を理論的に考察した。

ここでは、まず K-L 法を用いた教師なし類別法が

距離の概念を用いた類別方法であることを K-L 展開の基本的な立場から明らかにする<sup>3), 5)</sup>。つぎに、この類別方法を用いて主目的である教師なしの類別を行なう新しいアルゴリズムを明らかにする。最後に、先に筆者が発表した低次元の再構成したパターンを用いて行なったシミュレーション実験による類別結果を示す。

## 2. Karhunen-Lóeve 直交展開

Karhunen-Lóeve 展開は、与えられたパターンから特徴を抽出する最も有力な方法であるが、ここでは、後にこの展開を用いた類別方法が距離の概念を用いた類別方法であることを説明するため、まず Karhunen-Lóeve 展開の考えを導入する。この展開方法は確率ベクトルを直交ベクトル成分の線形結合で表示する方法である。いま、分布が未知なる正規化されたパターン  $X_i^{(k)}$  の集合  $S$  を

$$S = \{X_i^{(k)} | k=1 \sim K, i=1 \sim L\}. \quad (1)$$

ただし

$X_i^{(k)}$ :  $m$  次元パターンベクトル.

$$X^{(k)T} = (x_{i_1}^{(k)}, x_{i_2}^{(k)}, \dots, x_{i_m}^{(k)}).$$

$K$ : パターンクラスの数.

$L$ : 任意に抽出されたサンプルパターンの数.

$T$ : 転置したベクトルを表わす.

とすると、パターンベクトル  $X_i^{(k)}$  は、

$$X_i^{(k)} = \sum_{l=1}^m a_{il}^{(k)} \cdot \xi_l. \quad (2)$$

ただし、

† Unsupervised Clustering in an Orthogonal System, by Toyoshi Sera (Electrical Engineering Course Yamaguchi University Technical Junior College)

†† 山口大学工業短期大学部電気工学科

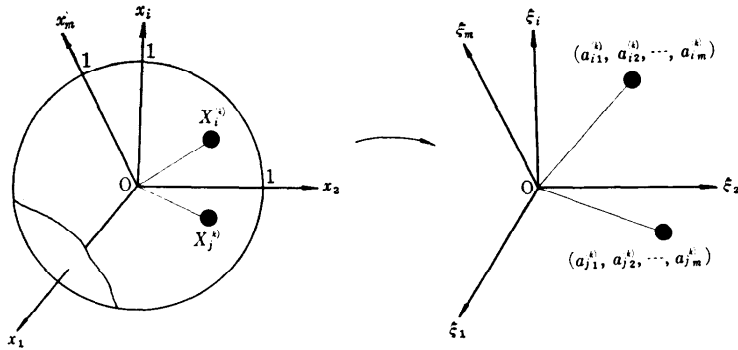


Fig. 1 Mapping two pattern vectors into the K-L orthogonal space

$\{\xi_l | l=1 \sim m\}$ : 正規直交ベクトル. (3)

$(\xi_i, \xi_j) = \delta_{ij}$ : クロネッカー  $\delta$ .

$\{(a_{il}^{(k)} | k=1 \sim K, i=1 \sim L, l=1 \sim m)\}$ :

展開係数. (4)

$a_{il}^{(k)} = (X_i^{(k)}, \xi_l)$

と展開されることがすでに明らかになっている<sup>2)</sup>. ここで  $\{a_{il}^{(k)}\}$  は、直交系  $\{\xi_l\}$  によって展開された各パターンの展開係数である。さて、実際に  $X_i^{(k)}$  を (2) 式の様に展開するためには直交系  $\{\xi_l\}$  の要素を求める必要がある。この要素は各クラスの生起確率を  $P_k$  とすれば多数のクラスがある場合でも、つぎの自己相関行列の固有ベクトルとして求められることが分っている<sup>4)</sup>。

$$G = \sum_{k=1}^K P_k \cdot E(X_i^{(k)} \cdot X_i^{(k)T}). \quad (5)$$

ただし、

$$\sum_{k=1}^K P_k = 1.$$

$E(X_i^{(k)} \cdot X_i^{(k)T})$ : クラス  $k$  に関する自己相関行列。

したがって、この対称行列  $G$  の固有値  $\lambda_l$  の集合を

$$\{\lambda_l | l=1 \sim m\} \quad (6)$$

とすれば、(3) 式の要素は

$$G \xi_l = \lambda_l \xi_l \quad (7)$$

から求められる。そして、この求められた  $\{\xi_l\}$  を、Karhunen-Lóeve 直交系、略して K-L 系という。さらに (7) 式から  $\lambda_l$  は

$$\lambda_l = \xi_l^T G \xi_l = \sum_{k=1}^K P_k \text{Var}(a_{il}^{(k)}). \quad (8)$$

$\text{Var}(a_{il}^{(k)})$ :  $a_{il}^{(k)}$  の分散。

と求められる。この結果、(4) 式と (8) 式から、 $\lambda_l$  は直交ベクトル  $\xi_l$  のパターンを表示する重要度すな

わち特徴量を表わしていることが分かる。このことは  $\{\xi_l\}$  の要素を座標軸とする空間にパターンベクトルが写像された状態を考えることによってより明らかになる。この様子を Fig. 1 に示す。

### 3. 教師なしの類別基準

2つの未知分布から抽出されたパターン集合を同一クラスに属するパターン同志に分ける2分類問題を考える。したがってパターン集合  $S$  は、

$$S = \{X_i^{(k)} | k=1, 2, i=1 \sim L\} \quad (9)$$

となる。さて、この集合を2つのクラスに教師なしで類別するには、教師の代りとなる一つの類別基準を定める必要がある。そこで最初、任意に分類された状態でのクラス1とクラス2のパターンの個数をそれぞれ  $M, N$  とし、さらに全パターンが (7) 式で求めた K-L 系  $\{\xi_l\}$  で表示されている (変換されている) ことを考慮すると、クラス1, クラス2の自己相関行列  $G_1, G_2$  は、クラス1, クラス2の各パターンが (4) 式の係数を用いて、

$$X_i^{(1)} = \sum_{l=1}^m a_{il}^{(1)} \xi_l, \quad X_i^{(2)} = \sum_{l=1}^m a_{il}^{(2)} \xi_l \quad (10)$$

と表わされることから、

$$G_1 = \frac{1}{M} \sum_{i=1}^M X_i^{(1)} \cdot X_i^{(1)T}. \quad (11)$$

$$G_2 = \frac{1}{N} \sum_{i=1}^N X_i^{(2)} \cdot X_i^{(2)T} \quad (12)$$

ただし、 $L = M + N$ : パターン総数。

となる。ここで (6) 式同様、自己相関行列  $G_1, G_2$  の固有値を

$$\left. \begin{aligned} \{\lambda_l^{(1)} | l=1 \sim m, \lambda_l^{(1)} = \frac{1}{M} \sum_{i=1}^M (a_{il}^{(1)})^2\} \\ \{\lambda_l^{(2)} | l=1 \sim m, \lambda_l^{(2)} = \frac{1}{N} \sum_{i=1}^N (a_{il}^{(2)})^2\} \end{aligned} \right\} \quad (13)$$

とし、類別基準  $C$  を距離の概念を用いて、

$$C = \sum_{i=1}^m (\lambda_i^{(1)} - \lambda_i^{(2)})^2 \quad (14)$$

と定義すれば、(4)式とさきに説明した(8)式の $\lambda_i$ のもつ意味、さらに Fig. 1 からこの基準  $C$  は、K-L 空間内でのクラス 1 とクラス 2 間の確率平均された特徴量(各パターンのもつ特徴量を確率平均したもの)の差の平方を表わしていることが分かる。したがって、全パターンが同じクラス同志にうまく類別されていれば、基準  $C$  の値は、そうでないときに比べてより大きな値をとることになるから、これより類別基準値を最大にするようにパターンを類別すれば、そのとき最適な結果が得られることが分かる。しかし、(13)式の要素  $\{\lambda_i^{(1)}\}$ ,  $\{\lambda_i^{(2)}\}$  を基準値  $C$  を改善するごとに求めることは大変な計算となる。そこで、自己相関行列  $G_1$  と  $G_2$  の差の行列を  $G_0$  とし、そのトレースを考えると、 $\{E_i\}$  系による変換前と変換後においてかわりはないので、

$$\begin{aligned} t_r(G_0) &= t_r(G_1 - G_2) \\ &= t_r \left( \frac{1}{M} \sum_{i=1}^M X_i^{(1)} X_i^{(1)T} - \frac{1}{N} \sum_{i=1}^N X_i^{(2)} X_i^{(2)T} \right) \\ &= t_r \begin{pmatrix} \left( \frac{1}{M} \sum_{i=1}^M (a_{i1}^{(1)})^2 \right) & & & 0 \\ -\frac{1}{N} \sum_{i=1}^N (a_{i1}^{(2)})^2 & & & \\ & \ddots & & \\ 0 & & \left( \frac{1}{M} \sum_{i=1}^M (a_{im}^{(1)})^2 \right) & \\ & & -\frac{1}{N} \sum_{i=1}^N (a_{im}^{(2)})^2 & \end{pmatrix} \\ &= t_r \begin{pmatrix} (\lambda_1^{(1)} - \lambda_1^{(2)}) & & & 0 \\ & (\lambda_2^{(1)} - \lambda_2^{(2)}) & & \\ & & \ddots & \\ 0 & & & (\lambda_m^{(1)} - \lambda_m^{(2)}) \end{pmatrix} \end{aligned} \quad (15)$$

となり、これより類別基準  $C$  は

$$\begin{aligned} C &= \sum_{i=1}^m (\lambda_i^{(1)} - \lambda_i^{(2)})^2 = \sum_{i=1}^m (\lambda_i^{(0)})^2 \\ &= t_r(G_0)^2. \end{aligned} \quad (16)$$

ただし、

$$\lambda_i^{(0)} = \lambda_i^{(1)} - \lambda_i^{(2)}: G_0 \text{ の固有値.}$$

と求められる。この結果  $G_0$  のトレースを最大にすることによってパターンが類別されることが分る。上式(16)は、先に富田らが明らかにした概念にもとずいて

導びかれたものであり、同じ結果を与えている。(詳しくは文献 5) を参照せよ。)

#### 4. 類別基準値 $C$ の最適化のアルゴリズム

教師なしの類別に必要な類別基準  $C$  は、Karhunen-Lóeve 直交系によって求められる特徴量に距離の概念を導入することによって、最終的に(16)式として求められた。しかし、実際にこの  $C$  を最適な値にもっていくためには、サンプルパターンを順次交換することが必要となる。両者のクラスのパターンを同時に1つずつ変換しながら  $C$  を最適にもって行くアルゴリズムはすでに明らかにされているが<sup>3)</sup>、このアルゴリズムは両者のクラスのパターン数が同じであるという条件、および1つ誤まって類別されると必然的に2つのパターンが誤まって類別されるという欠点をもっている。そこで、ここでは前半この両クラスのパターンを同時に交換するアルゴリズムを利用し、後半一方のクラスのパターンを他方のクラスに付加するアルゴリズムを利用する改良されたアルゴリズムを考察する。そのため、(9)式で与えられるパターンの集合が、任意にパターン数  $M$  個とパターン数  $N$  個のパターンクラスに分けられたとする。この状態を初期状態と呼ぶ。この結果、初期状態のクラス 1 の自己相関行列  $G_1$  とクラス 2 の自己相関行列  $G_2$  の差の行列  $G_0$  は

$$G_0 = \frac{1}{M} \sum_{i=1}^M X_i^{(1)} \cdot X_i^{(1)T} - \frac{1}{N} \sum_{i=1}^N X_i^{(2)} \cdot X_i^{(2)T} \quad (17)$$

ただし、

$$L = M + N: \text{総パターン数.}$$

となる。また、類別基準値  $C$  は(16)式から、

$$C = t_r(G_0)^2 \quad (18)$$

と求まる。さて、ここでの目的は、類別基準値  $C$  が最大になるようなクラスをクラス間でパターン交換、あるいは付加することによって求めるアルゴリズムを与えることである。そこで、まずクラス 1 の任意パターン  $X_k^{(1)}$  とクラス 2 の任意パターン  $X_l^{(2)}$  を交換する場合を考える。パターン交換の結果、クラス 1 とクラス 2 の自己相関行列  $G_1$ ,  $G_2$  は

$$\begin{aligned} G_1 &= \frac{1}{M} \sum_{i=1}^M X_i^{(1)} \cdot X_i^{(1)T} \\ &\quad - \left( \frac{1}{M} X_k^{(1)} \cdot X_k^{(1)T} - \frac{1}{M} X_l^{(2)} \cdot X_l^{(2)T} \right). \end{aligned} \quad (19)$$

$$G_2 = \frac{1}{N} \sum_{i=1}^N X_i^{(2)} \cdot X_i^{(2)T} - \left( \frac{1}{N} X_i^{(2)} \cdot X_i^{(2)T} - \frac{1}{N} X_i^{(1)} \cdot X_i^{(1)T} \right) \quad (20)$$

と書き改められるので、交換後の行列  $G_1$  と  $G_2$  の差の行列を  $G_0'$  とすれば (17) 式より、

$$G_0' = G_0 + \left( \frac{M+N}{MN} \right) \cdot \Delta G. \quad (21)$$

ただし、

$$\Delta G = (X_i^{(2)} \cdot X_i^{(2)T} - X_i^{(1)} \cdot X_i^{(1)T}) \quad (22)$$

となる。この結果、パターン交換による (20) 式の類別基準値  $C$  の変化を  $\Delta C$  とすれば、

$$\begin{aligned} \Delta C &= t_r(G_0')^2 - t_r(G_0)^2 \\ &= 2 \left( \frac{M+N}{MN} \right) t_r(G_0 \cdot \Delta G) + \left( \frac{M+N}{MN} \right)^2 t_r(\Delta G)^2 \end{aligned} \quad (23)$$

と求まる。したがって、 $\Delta C > 0$  なる場合は、交換の結果基準値  $C$  が  $\Delta C$  だけ増大することになり、パターンの交換以前より望ましいクラスが形成されることになるのでパターンの交換を行なう。 $\Delta C < 0$  の場合は逆になる。つぎに、クラス 1 のパターン  $X_i^{(1)}$  をクラスに付加する場合を考える。このときは行列  $G_1$ 、 $G_2$  は

$$\begin{aligned} G_1 &= \frac{M}{M-1} \left( \frac{1}{M} \sum_{i=1}^M X_i^{(1)} \cdot X_i^{(1)T} - \frac{1}{M} X_i^{(1)} \cdot X_i^{(1)T} \right), \quad (24) \\ G_2 &= \frac{N}{N+1} \left( \frac{1}{N} \sum_{i=1}^N X_i^{(2)} \cdot X_i^{(2)T} + \frac{1}{N} X_i^{(1)} \cdot X_i^{(1)T} \right) \end{aligned} \quad (25)$$

となり、差の行列  $G_0'$  は、

$$\begin{aligned} G_0' &= G_1 - G_2 \\ &= \frac{MN}{(M-1)(N+1)} \cdot G_0 + \Delta G_0. \end{aligned} \quad (26)$$

ただし、

$$\begin{aligned} \Delta G_0 &= \frac{1}{(M-1)(N+1)} \left\{ \left( \sum_{i=1}^M X_i^{(1)} \cdot X_i^{(1)T} + \sum_{i=1}^N X_i^{(2)} \cdot X_i^{(2)T} \right) - (M+N) X_i^{(1)} X_i^{(1)T} \right\} \end{aligned} \quad (27)$$

となる。したがって、 $M$ 、 $N$  が十分大きければ (23) 式と同様に類別基準値  $C$  の変化  $\Delta C$  は、

$$\begin{aligned} \Delta C &= t_r(G_0')^2 - t_r(G_0)^2 \\ &\approx 2t_r(G_0 \cdot \Delta G) + t_r(\Delta G)^2 \end{aligned} \quad (28)$$

と求まる。逆に、クラス 2 のパターン  $X_i^{(2)}$  をクラス 1 に付加した場合には、

$$G_0' = \frac{1}{(M+1)(N-1)} \cdot G_0 + \Delta G_0, \quad (29)$$

ただし、

$$\begin{aligned} \Delta G_0 &= \frac{1}{(M+1)(N-1)} \left\{ (M+N) X_i^{(2)} X_i^{(2)T} - \left( \sum_{i=1}^M X_i^{(1)} X_i^{(1)T} + \sum_{i=1}^N X_i^{(2)} X_i^{(2)T} \right) \right\} \end{aligned} \quad (30)$$

となり、類別基準値  $C$  の変化  $\Delta C$  は (28) 式で求められる。したがって、パターン交換の時と同様に、(28) 式の  $\Delta C$  が  $\Delta C > 0$  のときは、パターンを他方のクラスに付加することによってパターン付加以前よりも望ましいクラスが形成される。 $\Delta C < 0$  の場合は逆となる。ここで、(23) 式、(28) 式の第 2 項は常に正であることから  $\Delta C$  は、

$$\Delta C = 2 \left( \frac{M+N}{M \cdot N} \right) t_r(G_0 \cdot \Delta G), \quad (23)'$$

$$\Delta C = 2t_r(G_0 \cdot \Delta G) \quad (28')$$

とすることが出来る。以上により、類別を行なうアルゴリズムはつぎのようになる。

- 1) 与えられたパターン集合  $S$  を任意にパターン数  $M$  個と  $N$  個の 2 つのクラスに類別する。
- 2) クラス 1 内の任意パターン  $X_r^{(1)}$  の  $r$  ( $r=1, 2, \dots, M$ ) を  $r=1$  に固定し  $t$  ( $t=1, 2, \dots, N$ ) を 1 から  $N$  まで変化させ (23') 式の  $\Delta C$  を求める。
- 3)  $\Delta C > 0$  なるものの中で値が最大となるクラス 2 の  $t$  番目のパターン  $X_t^{(2)}$  をクラス 1 の  $X_r^{(1)}$  と交換する。
- 4)  $r=r+1$  として、 $r=M$  になるまで 2), 3) を繰り返す。
- 5) もはや  $\Delta C$  が  $\Delta C > 0$  となりえなくなったら、 $r$  を 1 から  $M$  まで変化させ (28') 式の  $\Delta C$  を求める。
- 6)  $\Delta C > 0$  なるものの中で値が最大となる  $r$  番目のパターン  $X_r^{(1)}$  をクラス 2 に付加する。すべての  $\Delta C$  が  $\Delta C \leq 0$  なら 8) に行く。
- 7) クラス 1 とクラス 2 のパターン数を  $M=M-1$ 、 $N=N+1$  と修正し 6) に帰る。
- 8) 6), 7) とは逆に、 $t$  を 1 から  $N$  まで変化させ (28') 式から  $\Delta C > 0$  なるものの中で値が最大となる  $t$  番目のパターン  $X_t^{(2)}$  をクラス 1 に付加する。すべての  $\Delta C$  が  $\Delta C \leq 0$  ならストップする。

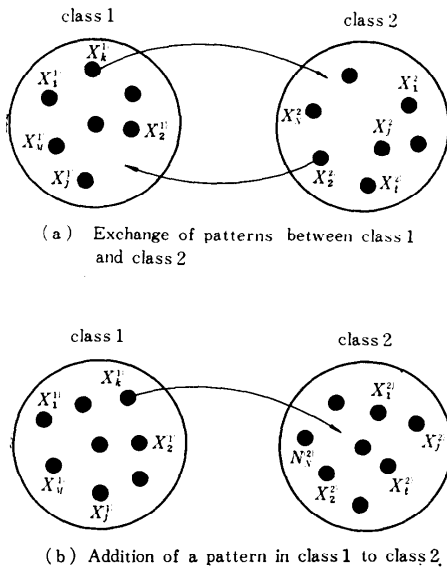


Fig. 2 Illustration of the exchange and addition of pattern

9) クラス1とクラス2のパターン数を  $M=M+1$ ,  $N=N-1$  として8)に帰る。

このアルゴリズムを遂行することにより、 $C$ の値の最大が求められ、その結果、最適な類別が教師なしに行なえる。Fig. 2に、パターンの交換とパターン付加の様子を図に示す。

5. 実験方法

さきに明らかにしたアルゴリズムによってパターン集合が類別出来ることを示すため、手書文字から再構成された低次元パターンの集合<sup>6)</sup>を用いてシミュレーション実験を行なった。パターンの集合としては、パターン間の従属割合(くわしくは文献6)を参照せよ。)が大きい場合と小さい場合の両者を採用した。アルゴリズム中の計算の処理は、最初パターン集合を任意にクラス1とクラス2に分け、(27), (30)式の

$$GT = \sum_{i=1}^M X_i^{(1)} \cdot X_i^{(1)T} + \sum_{i=1}^N X_i^{(2)} \cdot X_i^{(2)T} \quad (31)$$

各クラスの自己相関行列の差の行列  $G_0$ 、ならびに初期状態の基準値  $C$  を求めておく。以後はこれを用いて、1)~9)までの手順を行うことによって処理が行なわれる。Fig. 3にこのシミュレーション実験で用いた、第4節のアルゴリズムによる教師なし類別、処理プログラムの流れ図を示す。

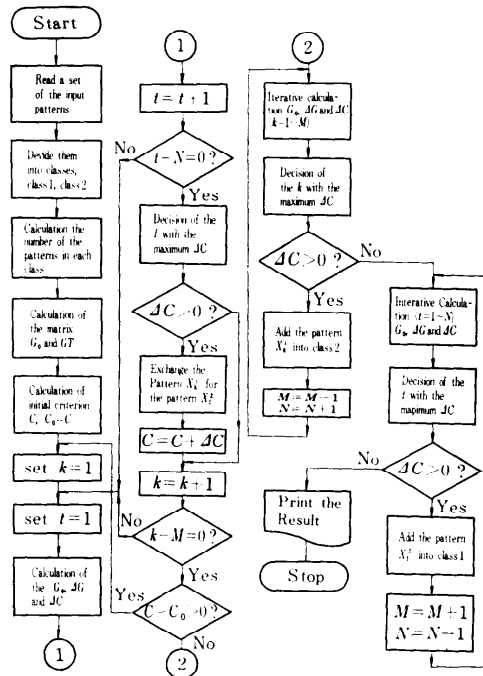


Fig. 3 Flow-Chart

6. 実験結果

実験は、手書文字  $A, B, E, F$  から再構成された低次元パターン(9次元)  $A, B$  と  $E, F$  の2組のパターンからなる約30個のパターン集合を2分類する実験を行なった。この際、

Case (a): 異なったクラスのパターン間の従属割合が大きいとき、

Case (b): 異なったクラスのパターン間の従属割合が小さいとき

に分けてそれぞれの組に対して実験を行なった。この結果、両者の場合に対して良い結果が得られた。Table 2にはCase (b)の場合のクラス1とクラス2の初期状態を  $A, B$  について示している。Table 2には類別結果を示している。Table 2から1つのパターン  $B$  が誤まって類別されていることが分るが、これは(7)式から求められた  $\{\xi_i\}$  空間上に  $A, B$  の全パターンをプロットした Fig. 4の分布の様子から誤まって類別されることが容易に理解出来る。この Fig. 4で、横軸、縦軸は  $\lambda_i$  の最も大きな値に対応する2つのベクトルを  $\xi_1, \xi_2$  として採用している。Fig. 5には、そのときの類別基準  $C$  の値が、計算処理中  $\Delta C > 0$

Table 1 Initial states of class 1 and class 2

		Class 1									Class 2												
Dimension	Pattern	1	2	3	4	5	6	7	8	9	N.f*	Dimension	Pattern	1	2	3	4	5	6	7	8	9	N.f*
A	A	0.0	6.0	2.0	2.0	5.0	4.0	3.0	0.0	2.0	98.0	A	A	1.0	3.0	1.0	3.0	4.0	4.0	2.0	1.0	2.0	61.0
B	B	1.0	2.0	2.0	2.0	4.0	1.0	4.0	1.0	2.0	51.0	B	B	1.0	0.0	1.0	2.0	2.0	1.0	1.0	1.0	0.0	13.0
B	A	3.0	2.0	1.0	3.0	3.0	1.0	2.0	2.0	2.0	45.0	A	A	0.0	4.0	1.0	2.0	3.0	3.0	1.0	0.0	2.0	44.0
A	A	0.0	4.0	3.0	1.0	4.0	5.0	3.0	0.0	1.0	77.0	A	A	0.0	2.0	2.0	2.0	2.0	4.0	2.0	0.0	1.0	37.0
A	B	0.0	1.0	3.0	1.0	3.0	4.0	2.0	0.0	2.0	44.0	A	A	1.0	2.0	1.0	2.0	3.0	4.0	2.0	1.0	1.0	41.0
B	B	2.0	1.0	1.0	2.0	1.0	1.0	2.0	0.0	0.0	16.0	B	B	3.0	2.0	1.0	4.0	2.0	1.0	3.0	1.0	2.0	49.0
B	B	3.0	2.0	1.0	4.0	2.0	2.0	3.0	1.0	1.0	49.0	B	B	1.0	0.0	0.0	1.0	3.0	1.0	3.0	0.0	2.0	25.0
B	B	1.0	1.0	1.0	2.0	2.0	1.0	2.0	2.0	1.0	21.0	A	A	1.0	3.0	1.0	3.0	4.0	4.0	2.0	1.0	2.0	61.0
A	A	1.0	2.0	3.0	3.0	4.0	5.0	2.0	0.0	2.0	72.0	A	A	0.0	3.4	1.0	1.0	4.0	4.0	3.0	1.0	2.0	57.0
B	B	2.0	1.0	1.0	2.0	1.0	1.0	2.0	0.0	0.0	16.0	A	A	0.0	2.0	1.0	0.0	3.0	4.0	3.0	1.0	1.0	41.0
B	B	2.0	1.0	2.0	3.0	2.0	1.0	2.0	1.0	0.0	28.0	B	B	2.0	1.0	0.0	3.0	3.0	1.0	3.0	1.0	3.0	43.0
B	B	3.0	2.0	1.0	4.0	2.0	1.0	3.0	1.0	2.0	49.0	A	A	0.0	3.0	1.0	1.0	4.0	4.0	3.0	1.0	2.0	57.0
B	B	0.0	1.0	1.0	1.0	2.0	1.0	1.0	1.0	0.0	10.0	A	A	0.0	2.0	2.0	1.0	3.0	3.0	2.0	0.0	2.0	35.0
A	A	1.0	2.0	1.0	2.0	3.0	4.0	2.0	1.0	1.0	41.0	B	B	3.0	2.0	1.0	4.0	2.0	2.0	3.0	1.0	1.0	49.0
A	A	0.0	3.0	1.0	1.0	4.0	4.0	3.0	1.0	2.0	57.0	A	A	0.0	4.0	3.0	1.0	4.0	4.0	3.0	0.0	1.0	68.0
B	B	3.0	2.0	1.0	4.0	1.0	1.0	4.0	2.0	2.0	56.0	A	A	0.0	3.0	2.0	2.0	3.0	3.0	1.0	0.0	2.0	40.0
A	A	0.0	2.0	1.0	0.0	3.0	4.0	3.0	1.0	1.0	41.0			0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0			0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0			0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0			0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0			0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

\* N.f: Normalized factor

Table 2 Final states of class 1 and class 2

		Class 1									Class 2												
Dimension	Pattern	1	2	3	4	5	6	7	8	9	N.f*	Dimension	Pattern	1	2	3	4	5	6	7	8	9	N.f*
B	B	3.0	2.0	1.0	4.0	2.0	1.0	0.3	1.0	2.0	49.0	A	A	0.0	3.0	1.0	1.0	4.0	4.0	3.0	1.0	2.0	57.0
B	B	3.0	2.0	1.0	4.0	2.0	1.0	3.0	1.0	2.0	49.0	A	A	0.0	4.0	3.0	1.0	4.0	5.0	3.0	0.0	1.0	77.0
B	B	3.0	2.0	1.0	3.0	3.0	1.0	2.0	2.0	2.0	45.0	A	A	0.0	4.0	1.0	2.0	3.0	3.0	1.0	0.0	2.0	44.0
B	B	1.0	0.0	1.0	2.0	2.0	1.0	1.0	1.0	0.0	13.0	A	A	0.0	2.0	2.0	2.0	2.0	4.0	2.0	0.0	1.0	37.0
B	B	2.0	1.0	0.0	3.0	3.0	1.0	3.0	1.0	3.0	43.0	A	A	1.0	2.0	1.0	2.0	3.0	4.0	2.0	1.0	1.0	41.0
B	B	2.0	1.0	1.0	2.0	1.0	1.0	2.0	0.0	0.0	16.0	A	A	0.0	6.0	2.0	2.0	5.0	4.0	3.0	0.0	2.0	98.0
B	B	3.0	2.0	1.0	4.0	2.0	2.0	3.0	1.0	1.0	49.0	A	A	1.0	2.0	1.0	2.0	3.0	4.0	2.0	1.0	1.0	41.0
B	B	1.0	1.0	1.0	2.0	2.0	1.0	2.0	2.0	1.0	21.0	A	A	0.0	2.0	1.0	0.0	3.0	4.0	3.0	1.0	1.0	41.0
B	B	1.0	2.0	2.0	2.0	4.0	1.0	4.0	1.0	2.0	51.0	A	A	0.0	3.0	1.0	1.0	4.0	4.0	3.0	1.0	2.0	57.0
B	B	2.0	1.0	1.0	2.0	1.0	1.0	2.0	0.0	0.0	16.0	A	A	0.0	2.0	1.0	0.0	3.0	4.0	3.0	1.0	1.0	41.0
B	B	2.0	1.0	2.0	3.0	2.0	1.0	2.0	1.0	0.0	28.0	A	A	0.0	1.0	3.0	1.0	3.0	4.0	2.0	0.0	2.0	44.0
B	B	3.0	2.0	1.0	4.0	2.0	1.0	3.0	1.0	0.0	49.0	A	A	0.0	3.0	1.0	1.0	4.0	4.0	3.0	1.0	2.0	57.0
B	B	1.0	0.0	0.0	1.0	3.0	1.0	3.0	0.0	2.0	25.0	A	A	0.0	2.0	2.0	1.0	4.0	3.0	2.0	0.0	2.0	35.0
B	B	3.0	2.0	1.0	4.0	1.0	1.0	0.0	4.0	2.0	56.0	A	A	1.0	2.0	3.0	3.0	3.0	5.0	2.0	0.0	2.0	72.0
		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0			0.0	4.0	3.0	1.0	4.0	4.0	3.0	0.0	1.0	68.0
		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0			0.0	3.0	2.0	2.0	3.0	3.0	1.0	0.0	2.0	40.0
		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0			1.0	3.0	1.0	3.0	4.0	4.0	2.0	1.0	2.0	61.0
		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0			1.0	3.0	1.0	3.0	4.0	4.0	2.0	1.0	2.0	61.0
		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0			0.0	1.0	1.0	1.0	2.0	1.0	1.0	1.0	0.0	10.0
		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0			0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

\* N.f: Normalized factor

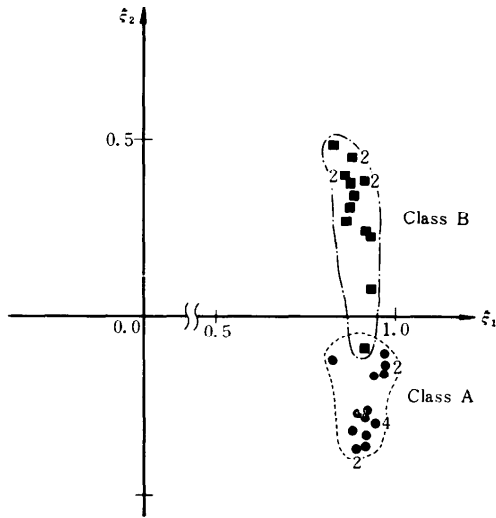


Fig. 4 Distribution of the patterns, A and B in Class (b)

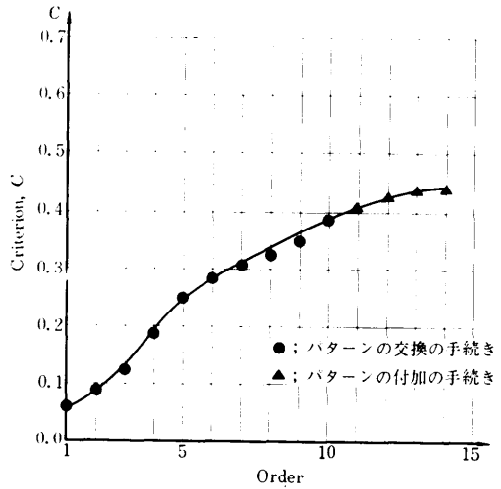


Fig. 5 Variation of Criterion, C

Table 3 Results of the classification

Case	Name	Initial State			Final State				Number of error Classification	
		Set of Patterns	Pattern	Class 1	Class 2	Set of Patterns	Pattern	Class 1		Class 2
(a)	A · B	A	A	10	6	A · B	A	16	0	0
		B	B	8	9		B	2	15	2
	E · F	E	E	10	6	E · F	E	12	4	4
		F	F	8	9		F	5	12	5
(b)	A · B	A	A	7	11	A · B	A	18	0	0
		B	B	10	5		B	1	14	1
	E · F	E	E	10	8	E · F	E	18	0	0
		F	F	6	11		F	0	17	0

で最大値をとる順番を横軸にプロットした様子を示す。この図中、丸印はパターンの交換による変化、三角印はパターンの付加による変化を示している。Table 3には、その他の場合の正しく類別された数と誤って類別された数を示している。そして、この表からクラスが異なるパターン同志の差が大きければうまく類別が出来るという当然の結果を示していることが分る。Table 4に、Case (b)でパターンの組がA, Bの場合の初期状態における差の行列  $G_0$ , (21), (26), (29)式から求められた最終状態における差の行列  $G_0$ , ならびに(31)式から求められた行列  $GT$  の計算結果をまとめて示す。

### 7. 結論

以上の考察ならびに実験結果より、

- (1) Karhunen-Löve 直交系を用いた類別法は類別基準  $C$  から分るように特徴量に距離の概念を導入したものである。これは富田らが明らかにした概念にもとづくものである。
- (2) そして、その最終の結果は2つのクラスの自己相関行列の差の行列  $G_0$  の平方のトレースとなる。
- (3) 考察された類別方法は、パターンの分布が未知でかつ分布が異なっている場合においても正し

Table 4 Result of the matrixes  $GT$  and  $G_0$

(I) Matrix  $GT$  in eq. (27)

2.110	1.048	1.044	2.812	2.055	1.435	2.425	0.8765	1.056
1.408	3.754	2.186	3.113	4.529	4.355	3.627	1.068	2.142
1.044	2.186	1.839	2.231	3.058	2.893	2.474	0.7275	1.254
2.812	3.113	2.231	4.625	4.396	3.569	4.125	1.511	2.127
2.055	4.529	3.058	4.396	6.757	5.931	5.295	1.729	3.062
1.432	4.355	2.893	3.569	5.931	6.192	4.558	1.284	2.645
2.425	3.627	2.474	4.125	5.295	4.558	4.997	1.530	2.545
0.8765	1.068	0.7275	1.511	1.729	1.284	1.530	0.8708	0.7552
1.056	2.142	1.254	2.127	3.062	2.654	2.545	0.7552	1.845

(II) Matrix  $G_0$  of initial state

0.1388	0.07285	0.05933	0.1714	0.1055	0.05342	0.1495	0.05248	0.05780
0.07285	-0.09881	-0.04489	0.02586	-0.1126	-0.1553	-0.02165	0.007902	-0.06822
0.05933	-0.04489	-0.02334	0.03833	-0.04668	-0.08749	0.004022	0.01685	-0.02524
0.1714	0.02586	0.03833	0.1831	0.06291	-0.01633	0.1478	0.06472	0.04025
0.1055	-0.1126	-0.04668	0.06291	-0.08356	-0.1828	0.008423	0.02008	-0.02827
0.05342	-0.1553	-0.08749	-0.01633	-0.1829	-0.2499	-0.08552	-0.01140	-0.07968
0.1495	-0.02165	0.004022	0.1478	0.008423	-0.08552	0.09697	0.03895	0.02723
0.05248	0.007902	0.01685	0.06472	0.02008	-0.01140	0.03895	0.02709	0.02065
0.05780	-0.03822	-0.02542	0.04045	-0.02287	-0.07968	0.02733	0.02065	0.009482

(III) Matrix  $G_0$  of final state

0.04719	0.02965	0.03034	0.04778	0.01630	0.01457	0.04345	0.01480	0.02300
0.02965	-0.02645	0.004572	0.007257	-0.01898	-0.04354	0.01496	0.01357	-0.02632
0.03034	0.004572	0.02188	0.02770	0.01889	0.001974	0.03403	0.01436	0.0005507
0.04778	0.007257	0.02770	0.03186	-0.006146	-0.01863	0.03926	0.01987	-0.01637
0.01630	-0.01898	0.01889	-0.006146	-0.03867	-0.04841	-0.005913	0.01407	-0.04203
0.01457	-0.04354	0.001974	-0.01863	-0.04841	-0.05584	-0.01431	-0.0005923	-0.03779
0.04345	0.01496	0.03403	0.03926	-0.005913	-0.01431	0.02959	0.02188	-0.02064
0.01480	0.01357	0.01436	0.01987	0.01407	-0.0005923	0.02188	0.01997	0.006152
0.02300	-0.02632	0.0005507	-0.01637	-0.04203	-0.03779	-0.02064	0.006152	-0.02953

くパターン集合を類別出来る方法である。

(4) パターンの交換手順にパターンを付加する手順を加えることによって、さきに Fukunaga<sup>3)</sup> が提案したパターン交換の手順の欠点を除くことが出来る。

ことが明らかになった。今後は、クラス間のパターン交換ならびに付加のアルゴリズムの改良を検討すると同時に対象パターンをかえた場合のシミュレーション実験を行なうつもりである。最後に、御世話になっている東京大学計数工学科甘利助教授ならびに山口大学工学部平田助教授に深く感謝します。同時に計算機で何かと手助けいただいた西村女史に感謝します。

参考文献

1) 飯島泰蔵: パターン認識の理論, 電気通信学会誌, Vol. 46, No. 11, pp. 1582~1590 (1963).

2) S. Watanabe: Knowing and Gessing, pp. 380~403, Wiley (1969).  
 3) K. Fukunaga and W. L. G. Koontz: Application of the Karhunen-Lóeve expansion to feature selection and ordering, IEEE Trans. C-19, pp. 311~318 (1970).  
 4) Y. T. Chein and K. S. Fu: Selection and Ordering of feature observations in a pattern recognition system, Information and Control 12, pp. 395~414 (1968).  
 5) 富田, 野口, 大泉: Karhunen-Lóeve 直交系による教師なしの類別, 電子通信学会誌, Vol. 54-C, No. 8, pp. 767~774 (1971).  
 6) 瀬良豊士: 3×3素子コードに Karhunen-Lóeve 直交系を用いたパターン認識, 情報処理, Vol. 13, No. 4, pp. 210~217 (1972).

(昭和48年6月11日受付)

(昭和48年7月24日再受付)