

# パターン付加の手続による教師なし類別

鳥 岡 豊 士\*

## Unsupervised Clustering with the Aid of the Adding-Procedure of a Pattern

Toyoshi TORIOKA

### Abstract

It is known that Karhunen-Lóeve orthogonal system can be applied to unsupervised clustering. And a method for unsupervised clustering by the K-L system were proposed and some results were given by the author. However, this method has two weak points as follow. (1) The algorithm of this method is complex. (2) Therefore, it is difficult for us to apply this method to multi-category problem.

Then, first the principle of the method for unsupervised clustering by K-L system are explained. Secondly, in order to get rid of these weak points, a explanation is given to a new method for unsupervised clustering by the use of the procedure which adds a pattern in one class to other class, where a part of this method has already shown in reference [8]. Finally, computer-simulated experiments are carried into effect in order to test usefulness of the present method.

### 1. 緒 言

普通の意味で、パターン認識を行なう場合、前もってクラスの集合が知られている。しかし、実際にこのクラスを形成するには、多数のパターン集合から共通の特徴を有するもの同志をまとめることによってクラスを形成する必要がある。すなわち、対象とするものを認知する必要がある。そこで、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 法を用いて、分布の未知なパターンからクラスを形成する類別問題 (Cognition 問題) を考察する。

K-L 系の基本的概念は、1963年飯島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 らは、この K-L 系を用いて、分散が等しい場合の 2 分類問題を考察し、教師なしの類別を提案した。さらに、1971年富田らは、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。1973年筆者は Fukunaga らの教師なし類別法の欠点を除いたパターン付加の手続を加えた新しい類別法を明らかにした。しかし、この方法もまだ二、三の欠点をもっている。そこで、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示す。つぎに、この類別方法を用いて、主目的である欠点を除いた教師なし、類別を行なうアルゴリズムとして、筆者が明らかにしたパターン付加の手続のみを用いた新しい方法を示す。最後に、実パターンを用いて行った実験結果を示す。

### 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  $X_i^{(k)}$  の集合を

$$S = \{ X_i^{(k)} \mid k = 1 \sim K, i = 1 \sim L \} \dots \dots \dots (1)$$

ただし

$X_i^{(k)}$  :  $m$  次元パターンベクトル

$$X_i^{(k)T} = (x_{i1}^{(k)}, x_{i2}^{(k)}, \dots, x_{im}^{(k)})$$

- $K$  : 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,

$$X_i^{(k)} = \sum_{\ell=1}^m a_{i\ell}^{(k)} \cdot \xi_{\ell} \quad \dots\dots\dots (2)$$

ただし

$$\{\xi_{\ell} \mid \ell = 1 \sim m\} : \text{正規直交ベクトル} \quad \dots\dots\dots (3)$$

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

$$\left\{ a_{i\ell}^{(k)} \mid k = 1 \sim K, i = 1 \sim L, \ell = 1 \sim m \right\} : \text{展開係数} \quad \dots\dots\dots (4)$$

$$a_{i\ell}^{(k)} = (X_i^{(k)} \cdot \xi_{\ell})$$

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  $\{a_{i\ell}^{(k)}\}$  は, 直交系  $\{\xi_{\ell}\}$  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) 式の様子に展開するためには直交系  $\{\xi_{\ell}\}$  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>.

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

ただし

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

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

したがって, この対称行列  $G$  の固有値  $\lambda_{\ell}$  の集合を  $\{\lambda_{\ell} \mid \ell = 1 \sim m\}$   $\dots\dots\dots (6)$

とすれば, (3) 式の要素は

$$G \cdot \xi_{\ell} = \lambda_{\ell} \xi_{\ell} \quad \dots\dots\dots (7)$$

から求められる. そして, この求められた  $\{\xi_{\ell}\}$  を, Karhunen-L6eve 直交系, 略して K-L 系という. さらに (7) 式から  $\lambda_{\ell}$  は

$$\lambda_{\ell} = \xi_{\ell}^T \cdot G \cdot \xi_{\ell} = \sum_{k=1}^K P_k \cdot \text{Var}(a_{i\ell}^{(k)}) \quad \dots\dots\dots (8)$$

$$\text{Var}(a_{i\ell}^{(k)}) : a_{i\ell}^{(k)} \text{ の分散}$$

と求められる. この結果, (4) 式と (8) 式から,  $\lambda_{\ell}$  は直交ベクトル  $\xi_{\ell}$  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,  $\{\xi_{\ell}\}$  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り明らかになる. この様子を Fig. 1 に示す.

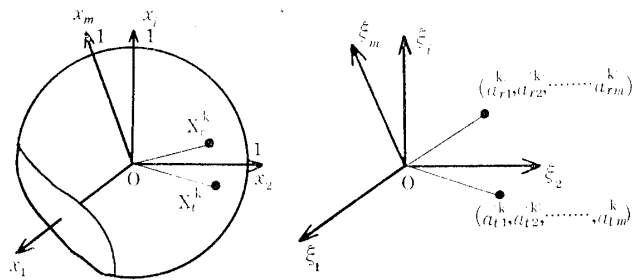


Fig. 1 Mapping two pattern vectors into the K-L orthogonal space

### 3. 教師なし類別を行うための類別基準

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,

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

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  $\{\xi_{\ell}\}$  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

$$X_i^{(1)} = \sum_{\ell=1}^m a_{i\ell}^{(1)} \xi_{\ell} \quad X_i^{(2)} = \sum_{\ell=1}^m a_{i\ell}^{(2)} \cdot \xi_{\ell} \quad \dots\dots\dots (10)$$

と表わされることから

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

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

ただし,

$$L = M + N : \text{パターン総数}$$

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

$$\left\{ \lambda_{\ell}^{(1)} \mid \ell = 1 \sim m, \lambda_{\ell}^{(1)} = \frac{1}{M} \cdot \sum_{i=1}^M (a_{i\ell}^{(1)})^2 \right\} \quad (13)$$

$$\left\{ \lambda_{\ell}^{(2)} \mid \ell = 1 \sim m, \lambda_{\ell}^{(2)} = \frac{1}{N} \cdot \sum_{i=1}^N (a_{i\ell}^{(2)})^2 \right\}$$

とし, 類別基準  $C$  を距離の概念を用い,

$$C = \sum_{\ell=1}^m (\lambda_{\ell}^{(1)} - \lambda_{\ell}^{(2)})^2 \quad \dots\dots\dots (14)$$

と定義すれば, (4) 式とさきに説明した (8) 式の  $\lambda_{\ell}$  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取るようになるから、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を改善するこ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、 $\{\xi_i\}$  系による変換前と変換後においてかわりないので、

$$\begin{aligned}
 \text{tr}(G_0) &= \text{tr}(G_1 - G_2) \\
 &= \text{tr}\left(\frac{1}{M} \cdot \sum_{i=1}^M X_i^{(1)} \cdot X_i^{(1)T} - \frac{1}{N} \cdot \sum_{i=1}^N X_i^{(2)} \cdot X_i^{(2)T}\right) \\
 &= \text{tr} \left[ \begin{array}{ccc} \left(\frac{1}{M} \cdot \sum_{i=1}^M (a_{i1}^{(1)})^2 - \frac{1}{N} \cdot \sum_{i=1}^N (a_{i1}^{(2)})^2\right) & & 0 \\ & 0 & \\ & & \left(\frac{1}{M} \cdot \sum_{i=1}^M (a_{im}^{(1)})^2 - \frac{1}{N} \cdot \sum_{i=1}^N (a_{im}^{(2)})^2\right) \end{array} \right] \\
 &= \text{tr} \left[ \begin{array}{ccc} \left(\lambda_1^{(1)} - \lambda_1^{(2)}\right) & & \\ & \left(\lambda_2^{(1)} - \lambda_2^{(2)}\right) & 0 \\ & & 0 \\ & & \left(\lambda_m^{(1)} - \lambda_m^{(2)}\right) \end{array} \right] \dots\dots\dots (15)
 \end{aligned}$$

となり、これより類別基準Cは

$$\begin{aligned}
 C &= \sum_{i=1}^m \left(\lambda_i^{(1)} - \lambda_i^{(2)}\right)^2 = \sum_{i=1}^m \left(\lambda_i^{(0)}\right)^2 \\
 &= \text{tr}(G_0)^2 \dots\dots\dots (16)
 \end{aligned}$$

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づいて導かれるものであり、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最適化のアルゴリズム

教師なしの類別に必要な類別基Cは、Karhunen-Lóve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、実際にこの類別基準値Cを最適な値

にもっていくためにはパターンを順次交換する必要がある。この交換方法について、著者によりすでに明らかにされているが<sup>7),8)</sup>、ここではパターン付加のアルゴリズムのみを用いても類別基準値Cを最適な値にもっていくことが出来ることを示す。そのため、(9)式で与えられるパターン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

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

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

$$C = \text{tr}(G_0)^2 \dots\dots\dots (18)$$

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を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  $X_r^{(1)}$  をクラス2に付加する場合を考える。このパターン  $X_r^{(1)}$  の付加の結果のクラス1とクラス2の自己相関行列を  $G_1'$ ,  $G_2'$  とすれば(11), (12)式より

$$G_1' = \frac{M}{M-1} \cdot \left(G_1 - \frac{1}{M} \cdot X_r^{(1)} \cdot X_r^{(1)T}\right) \dots\dots\dots (19)$$

$$G_2' = \frac{N}{N+1} \cdot \left(G_2 + \frac{1}{N} \cdot X_r^{(1)} \cdot X_r^{(1)T}\right) \dots\dots\dots (20)$$

となるので、パターン付加後の自己相関行列  $G_1'$ ,  $G_2'$  の差の行列  $G_0'$  は(17), (19), (20)式より

$$\begin{aligned}
 G_0' &= G_1' - G_2' \\
 &= \frac{N \cdot M}{(N+1)(M-1)} \cdot G_0 + \frac{M+N}{(N+1)(M-1)} \cdot \\
 &\quad \left\{ M \cdot G_1 + N \cdot G_2 - (M+N) \cdot X_r^{(1)} \cdot X_r^{(1)T} \right\} \\
 &= \alpha G_0 + \beta \cdot \Delta G_0 \dots\dots\dots (21)
 \end{aligned}$$

ただし、

$$\alpha = \frac{N \cdot M}{(N+1) \cdot (M-1)}$$

$$\beta = \frac{1}{(N+1) \cdot (M-1)}$$

$$\Delta G_0 = L \cdot G_1 + M \cdot G_2 - (M+L) \cdot X_r^{(1)} \cdot X_r^{(1)T}$$

となる。したがって、パターン付加による類別基準値の変化を  $\Delta C$  とすれば

$$\begin{aligned} \Delta C &= t_r(G_0)^2 - t_r(G_0)^2 \\ &= t_r\{(\alpha \cdot G_0 - \beta \cdot \Delta G_0)^2 - G_0^2\} \\ &= (\alpha^2 - 1) \cdot t_r(G_0)^2 + 2\alpha \cdot \beta \cdot t_r(G_0 \cdot \Delta G_0) \\ &\quad + \beta^2 \cdot t_r(\Delta G_0)^2 \end{aligned} \dots\dots\dots (22)$$

となる。したがって、 $\Delta C > 0$  なる場合は、パターン付加の結果基準値  $C$  が  $\Delta C$  増大することになり、パターンの付加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付加を行なう。 $\Delta C < 0$  の場合は逆になるのでパターン付加を行なわない。つぎに、クラス2の任意パターン  $X_t^{(2)}$  をクラス1に付加する場合を考えると

$$G_1' = \frac{M}{M+1} \cdot \left( G_1 + \frac{1}{M} \cdot X_t^{(2)} X_t^{(2)T} \right) \dots\dots\dots (23)$$

$$G_2' = \frac{N}{N-1} \cdot \left( G_2 + \frac{1}{N} \cdot X_t^{(2)} X_t^{(2)T} \right) \dots\dots\dots (24)$$

となり、自己相関行列の差の行列  $G_0'$  は (17), (23), (24) 式より

$$\begin{aligned} G_0' &= G_1' - G_2' \\ &= \frac{MN}{(M+1)(N-1)} \cdot G_0 - \frac{1}{(M+1)(N-1)} \cdot \\ &\quad \left\{ L \cdot G_1 + M \cdot G_2 - (L+M) \cdot X_r^{(2)} X_r^{(2)T} \right\} \\ &= \alpha \cdot G_0 - \beta \cdot \Delta G_0 \end{aligned} \dots\dots\dots (25)$$

ただし

$$\alpha = \frac{M \cdot N}{(M+1) \cdot (N-1)}$$

$$\beta = \frac{1}{(M+1) \cdot (N-1)}$$

ともとまる。したがって、パターン付加による類別基準値  $\Delta C$  の変化は(22)と同様にして

$$\begin{aligned} \Delta C &= t_r(G_0')^2 - t_r(G_0)^2 \\ &= (\alpha^2 - 1) \cdot t_r(G_0)^2 \\ &\quad - 2 \cdot \alpha \cdot \beta \cdot t_r(G_0 \cdot \Delta G_0) + \beta^2 \cdot t_r(\Delta G_0)^2 \end{aligned} \dots\dots\dots (26)$$

ともとまる。したがって、 $\Delta C > 0$  のときは、パターン  $X_t^{(2)}$  をクラス1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。

$\Delta C < 0$  の場合には逆となる。ここで、(22), (26)式の第3項は

$$\beta^2 \cdot t_r(\Delta G_0)^2 \geq 0 \dots\dots\dots (27)$$

ゆえ

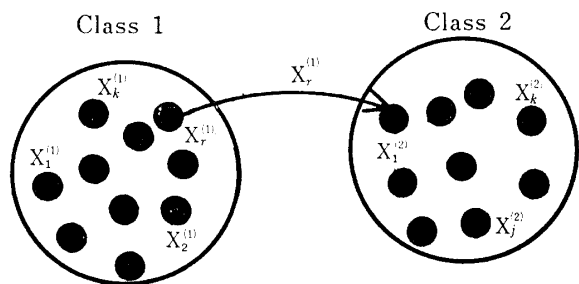
$$\begin{aligned} \Delta C &= (\alpha^2 - 1) \cdot t_r(G_0)^2 \\ &\quad + 2 \cdot \alpha \cdot \beta \cdot t_r(G_0 \cdot \Delta G_0) \end{aligned} \dots\dots\dots (22')$$

$$\begin{aligned} \Delta C &= (\alpha^2 - 1) \cdot t_r(G_0)^2 \\ &\quad - 2 \cdot \alpha \cdot \beta \cdot t_r(G_0 \cdot \Delta G_0) \end{aligned} \dots\dots\dots (26')$$

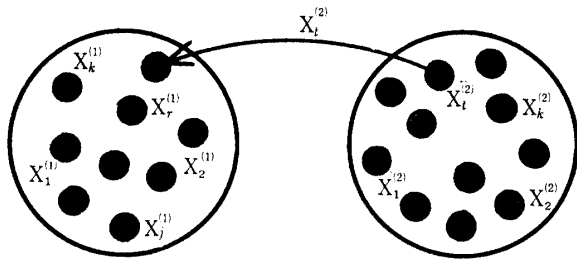
とすることが出来る。以上より、類別を行なうパターン付加のアルゴリズムはつぎのようになる。

- 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  $r=1$  から順次(22')式の  $\Delta C$  を計算し、 $\Delta C > 0$  なら  $X_r^{(1)}$  をクラス2に付加し、 $M=M-1, N=N+1, r=r-1$  とし3)に行く。 $\Delta C < 0$  なら  $r=r+1$  とし、 $t=M$  になるまで繰り返し、 $t=M$  になったら4)に行く。
- 3) クラス1内のパターンを並びかえ(付加したパターンの空手の所へ以下のパターンを繰り上げてつめる)、 $r=r+1$  とし  $r=M$  になるまで2)を繰返す。 $r=M$  になったら4)に行く。
- 4) クラス2内の任意パターン  $X_t^{(2)}$  の  $t(t=1, 2, \dots, N)$  を  $t=1$  から順次(26')式の  $\Delta C$  を計算し、 $\Delta C > 0$  なら  $X_t^{(2)}$  をクラス1に付加し、 $N=N-1, M=M+1, t=t-1$  とし4)に行く。 $\Delta C < 0$  なら  $t=t+1$  とし  $t=N$  になるまで繰返す。 $t=N$  なら5)に行く。
- 5) クラス内のパターンを並びかえ、 $t=t+1$  とし  $t=N$  になるまで繰返す。 $t=M$  になったら6)に行く。
- 6) 1) ~ 5) までのパターン付加による類別基準値の変化  $\Delta C > 0$  の総和が正であるなら、1) ~ 5) を繰返す。正となり得なくなったらストップする。

以上のアルゴリズムを遂行することにより、類別基準値  $C$  の最大が求められ、その結果として最適な類別が教師なしに行なわれる。Fig. 2 にパターン付加の様子を示す。



(a) Addition of a pattern from class1 to class2



(b) Addition of a pattern from class2 to class1

Fig. 2 Illustration for the addition of pattern

5. 実験方法

さきに明らかにしたパターン付加の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集合<sup>6)</sup>と5次元の正規分布から抽出されたパターン集合を用いて、シミュレーション実験を行なった。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(21), (25)式の行列

$$GT = M \cdot G_1 + N \cdot G_2 \quad \dots\dots\dots (28)$$

ならびに、各クラスの自己相関行列  $G_0$ 、ならびに初期状態の類別基準値  $C$  を求めておく。以後は、これらを用いて、(21), (25), (22'), (26)'式を用いて、1)~6) までのパターン付加のアルゴリズムを遂行することによって教師なしの類別が行なわれる。Fig. 3 に、このシミュレーション実験で用いた第4節のパターン付加による教師なし類別アルゴリズム処理プログラムの流れ図を示す。

6. 実験結果

実験は、まず実パターンである手書文字 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

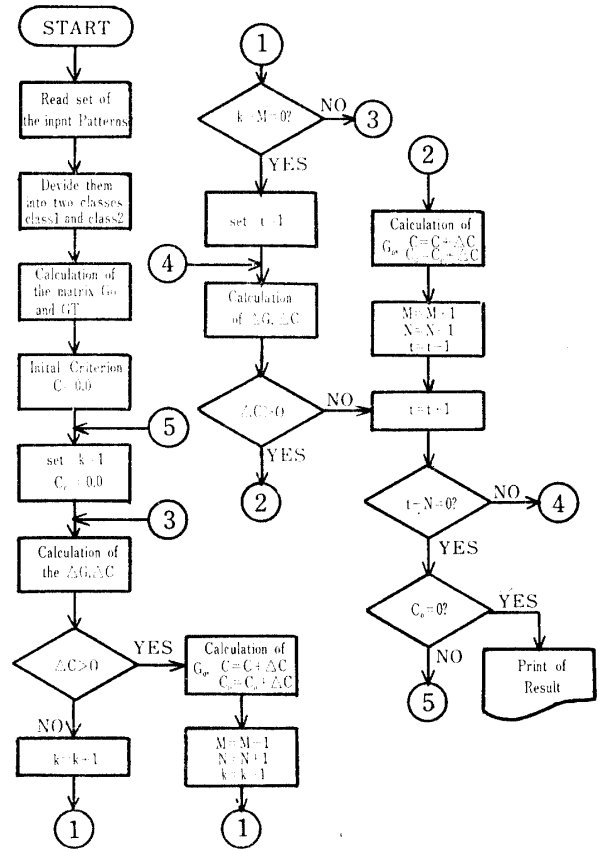


Fig. 3 Flow chart

ン集合を2分類する実験を行なった。この際

Case (a) : 集合 S がパターン A と B から成るとき。

Case (b) : 集合 S がパターン E と F から成るときに分けて実験を行なった。Table 1 に Case (a) と Case (b) の結果を示している。この表から両者の場合に対して良い結果が得られていることが分る。Fig. 4 には、Case (b) での類別基準値  $\Delta C > 0$  の和の変化を初期状態の類別基準値  $C$  を零とし、 $\Delta C > 0$  が起った順にプロットしている。この図中、角印はクラス1からクラス2へのパターン付加。三角印はクラス2

Table 1 Results of the Classification

State			Initial State		Final State		the number of error classification
case, Pattern Name			Class 1	Class 2	Class 1	Class 2	
Case	Set of Patterns	Pattern					
(a)	A · B	A	7	11	18	0	0
		B	9	5	1	13	1
(b)	F · E	E	10	8	17	0	0
		F	6	11	1	16	1

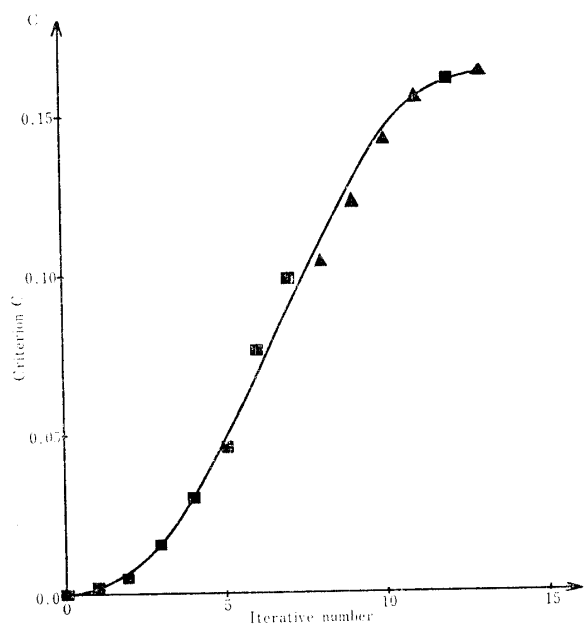


Fig. 4 Variation of Criterion C

からクラス1へのパターン付加による変化を示している。図5には、Case (b) のパターン集合 E, F の分布を示す。この図で座標軸 ( $\xi_1, \xi_2$ ) は、(5)式から

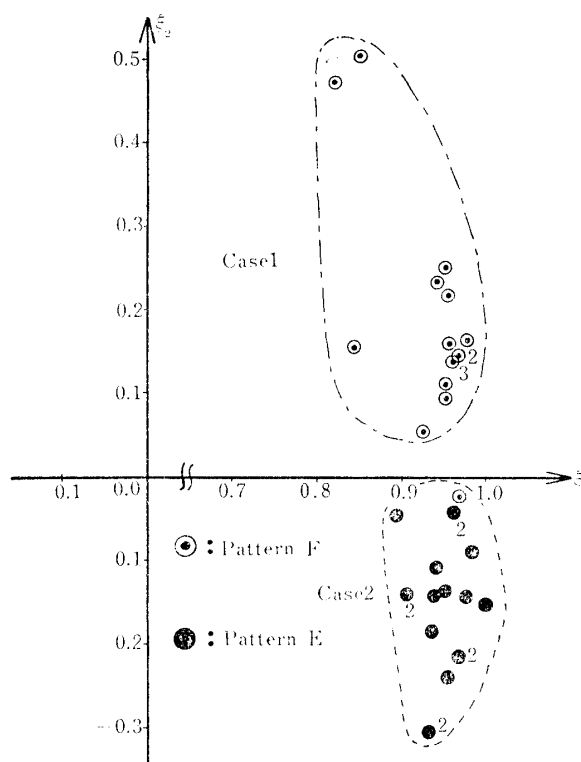


Fig. 5 Distribution of the patterns, E and F of Case(b)

Table 2 Result of the matrixes GT and G<sub>0</sub>

(I) Matrix GT								
11.7513	9.6792	3.4066	5.4079	3.6633	4.9144	5.7076	2.1926	1.4442
9.6792	8.5981	2.8091	4.4428	3.0861	3.9845	4.8439	1.9268	1.2635
3.4066	2.8069	1.7834	1.5534	1.4929	1.7951	1.7653	0.5185	0.2832
5.4079	4.4428	1.5534	3.0953	1.8982	2.3079	2.8692	1.1640	0.8679
3.6633	3.0861	1.4929	1.8982	1.8850	1.9432	2.0988	0.7512	0.6201
4.9144	3.9845	1.7951	2.3079	1.9432	2.5542	2.4688	0.8303	0.5411
5.7076	4.8439	1.7653	2.8692	2.0988	2.4688	3.1231	1.1237	0.9027
2.1926	1.9268	0.5185	1.1640	0.7512	0.8303	1.1237	0.7443	0.4450
1.4424	1.2635	0.2832	0.8679	0.6201	0.5411	0.9027	0.4450	0.4931
(II) Matrix G <sub>0</sub> of initial state								
+0.0172	-0.0175	+0.0239	-0.0361	+0.0141	+0.0452	-0.0021	-0.0154	-0.0270
-0.0175	-0.0331	+0.0043	-0.0497	+0.0031	+0.0256	-0.0109	-0.0199	-0.0231
+0.0239	+0.0043	+0.0125	+0.0194	+0.0253	+0.0308	+0.0212	+0.0008	+0.0028
-0.0361	-0.0497	+0.0194	-0.0518	+0.0348	+0.0099	-0.0186	-0.0213	-0.0201
+0.0141	+0.0031	+0.0253	+0.0035	+0.0378	+0.0300	+0.0199	-0.0010	+0.0055
+0.0452	+0.0256	+0.0308	+0.0099	+0.0300	+0.0419	+0.0240	-0.00179	-0.0040
-0.0022	-0.0109	+0.0212	-0.0186	+0.0199	+0.0240	+0.0048	-0.0077	-0.0091
-0.0154	-0.0199	+0.0008	-0.0213	-0.0010	-0.0018	-0.0077	-0.0094	-0.0070
-0.0270	-0.0231	+0.0028	-0.0201	+0.0055	-0.0040	-0.0091	-0.0070	-0.0048

(III) Matrix  $G_0$  of final state

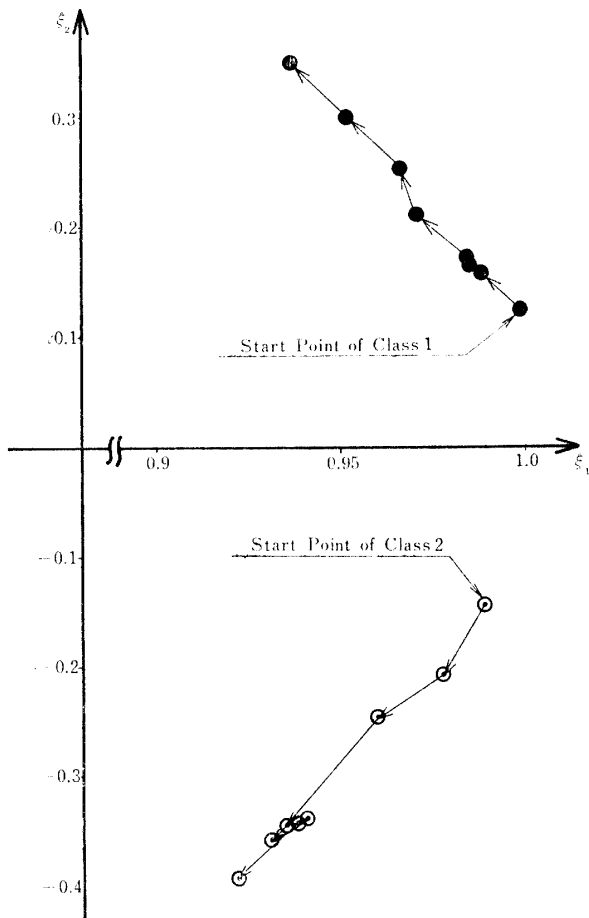
-0.0419	-0.0673	0.7342	-0.0411	0.0579	0.0856	-0.0336	-0.0634	-0.0525
-0.0673	-0.1022	0.0986	-0.0534	0.0255	0.0458	-0.0516	-0.0625	-0.0505
0.1342	0.0986	0.0920	0.0623	0.0760	0.0906	0.0666	0.0123	0.0080
-0.0411	-0.0534	0.0623	-0.0420	0.0252	0.0338	-0.0240	-0.0349	-0.0285
0.0579	0.0255	0.0760	0.0252	0.0565	0.0644	0.0261	0.0021	-0.0068
0.0856	0.0458	0.0906	0.0338	0.0644	0.0843	0.0370	-0.0102	-0.0103
-0.0336	-0.0516	0.0666	-0.0240	0.0261	0.0370	-0.0195	-0.0320	-0.0265
-0.0634	-0.0625	0.0123	-0.0349	0.0021	-0.0102	-0.0320	-0.0221	-0.0168
-0.0525	-0.0505	0.0080	-0.0285	-0.0068	-0.0103	-0.0265	-0.0168	-0.0156

求めた固有値で大きいものの2つに対応する固有ベクトルである。Table 2 に、Case (b) の場合の初期状態における差の行列、最終状態における差の行列  $G_0$ 、および(29)式の行列  $GT$  の結果をまとめて示している。

つぎに、教師なし類別を行なうにつれクラス間の距離がどのように変化していくかを知るため、つぎのような5次元正規分布から抽出された20個のパターン集合を用いて実験を行なった。

$$f(X^{(1)T}) = \frac{1}{(\sqrt{2\pi}\sigma)^5} \exp\left[-\frac{\{(x_1-4)^2-(x_2-5)^2+(x_3-2)^2+(x_4-3)^2+(x_5-2)^2\}}{2\sigma^2}\right] \dots\dots\dots (29)$$

$$f(X^{(2)T}) = \frac{1}{(\sqrt{2\pi}\sigma)^5} \exp\left[-\frac{\{(x_1-3)^2+(x_2-2)^2+(x_3-5)^2+(x_4-2)^2+(x_5-5)^2\}}{2\sigma^2}\right] \dots\dots\dots (30)$$



ただし

$$f(X^{(k)T}) : X^{(k)T} = (x_1, x_2, \dots, x_5) \text{ の密度関数.}$$

$$\sigma^2 : \sigma^2 = 1$$

Fig. 6 に距離の変化していく様子を示す。この図から  $\Delta C \geq 0$  になるごとに距離が大きくなって行く様子が良く分る。ただし、図中の各点は  $\Delta C > 0$  になったときの各クラスの最大固有値に対応する固有ベクトルを(5)式で求めたそのうちで、大きいもの2つに対する固有ベクトルを座標軸 ( $\xi_1, \xi_2$ ) とする空間にプロットしたものである。

### 7. 結 論

以上の考察ならびに実験結果より

- (1) さきに筆者が明らかにした教師なし類別アルゴリズムから、パターン交換の手続きを除いたパターン付加のみのアルゴリズムによっても十分教師なしの類別が行なわれる。
- (2) 先に明らかにしたアルゴリズムに比較して、 $\Delta C > 0$  である場合には、常にパターンを付加するようにアルゴリズムを作っているので早く処理が行なわれる。
- (3) このパターン付加のアルゴリズムは、先にのべた方法に比べて、アルゴリズムが簡単であるので、多分類教師なしアルゴリズムに簡単に拡張出

Fig. 6 Variation of distance between class1, and class 2

来る利点をもっている。

- (4) しかし、パターン付加がクラス1からクラス2、クラス2からクラス1へと行なわれるので、初期状態において、クラス1のクラス2の記憶容量が先に明らかにしたパターンに比較して多少多く必要とされる欠点がある。

ことが明らかになった。今後は、パターン集合  $S$  が多数のクラスのパターンから成っている場合の教師なし分類の問題について、検討すると同時に誤り分類の割合等についても考察を加えていくつもりである。最後に、計算機で何かと手助けいただいた西村女史に深く感謝します。

#### 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，46，1582—1590 (1963)
- 2) S. Watanabe：“Knowing and Guessing”，Wiley

(1969) P. 380—403

- 3) K. Fukunaga and W. L. Koontz：“Application of the Karhunen-Loève expansion to feature selection and ordering”，IEEE Trans. C—19, 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, 395—414 (1968)
- 5) 富田，野口，大泉：Karhunen-Loève 直交系による教師なしの類別，電子通信学会誌，54—C, 767—774 (1971)
- 6) 瀬良豊士：3×3素子コードに Karhunen-Loève 直交系を用いたパターン認識，情報処理，13, 210—217 (1972)
- 7) 瀬良豊士：“直交空間での教師なし類別について基礎研究”，山口大学工学部研究報告，24, 1—6 (1973)
- 8) 瀬良豊士：“直交空間の教師なし類別について”，情報処理，14, 746—753 (1973)

(昭和49年4月15日受理)