

# 主成分分析法によるサンプル次元の減少と クラスターリングについて

岡田 敏彦\*・富田 眞吾\*

## On Dimensionality Reduction of Pattern Samples by Principal Components Analysis and Clustering

Toshihiko OKADA and Shingo TOMITA

### Abstract

Several methods for dimensionality reduction in the clustering procedures have been proposed. This paper reviews that a method of dimensionality reduction by principal components analysis is formed linear combinations of the features of given patterns. The main object of principal components analysis is to find a lower-dimensional representation that accounts the variance of the features. Moreover, the method of iterative improvement to minimize the sum-of-squared-error criterion for clustering is described. And to show the usefulness of the above mentioned method, the simulation for concrete pattern samples on a computer is executed, and we can obtain the good agreement with our theoretical results.

### 1. ま え が き

クラスターリング技法についてはこれまでに種々の方法が提案されてきたが(文献1), 2)に詳しい) サンプルが高次元空間内に表示されたパターンベクトルの場合, クラスターリング過程での処理時間とかサンプルの記憶領域は多大なものとなり, したがって処理が不可能なような問題が生じる. もし高次元のサンプルからいくつかの有用な特徴量を抽出できるなら, それらの特徴量でもって, サンプルを記述することができ, これらの問題点は解決される. 本論文ではクラスターリングの前処理として, サンプルの特徴抽出を主成分分析法の概念を基礎にして行ない, 高次元サンプルを少数の特徴量で記述することにより, 低次元のサンプルに減じる方法を提案し, さらに低次元に減じられたサンプル集合のクラスターリングに関し, 教師なしで行なう一つの反復アルゴリズムについて述べている. さらに, いくつかの実験結果について報告し, 本方法の有用性について検討している.

### 2. サンプル行列, 重みベクトル, 合成変量ベクトル

本論文において用いられるサンプル行列, 重みベクトル, 合成変量ベクトルについて説明する.

$n$ 次元のサンプルが  $N$ 個与えられているとしよう. そのサンプルをまとめてサンプル行列の形で表わす.

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & & \vdots \\ x_{N1} & x_{N2} & \cdots & x_{Nn} \end{bmatrix} \quad (1)$$

行列  $X$  の各成分  $x_{ij}$  はサンプル  $i$  の変量  $j$  の測定値を表わす. 変量  $j$  の平均値を  $m_j$ , 分散を  $s_j^2$  とすると,

$$m_j = \left( \sum_{i=1}^N x_{ij} \right) / N \quad (j = 1, 2, \dots, n) \quad (2)$$

$$s_j^2 = \sum_{i=1}^N (x_{ij} - m_j)^2 / N \quad (3)$$

で与えられる. この  $m_j$  と  $s_j$  を使って行列  $X$  のすべての成分を列について標準化してできる行列を

$$Z = \begin{bmatrix} z_{11} & z_{12} & \cdots & z_{1n} \\ z_{21} & z_{22} & \cdots & z_{2n} \\ \vdots & \vdots & & \vdots \\ z_{N1} & z_{N2} & \cdots & z_{Nn} \end{bmatrix} \quad (4)$$

として表わす。ここに、 $z_{ij} = (x_{ij} - m_j) / s_j$ 。これは標準化されたサンプル行列で以下この  $Z$  をサンプル行列と呼ぶことにする。 $Z$  の各列は、すなわち各変量の平均値はいずれも 0 であり、分散は 1 となっているので、変量  $j$  と変量  $k$  の間の相関係数は

$$r_{jk} = (\sum_{i=1}^N z_{ij} z_{ik}) / N$$

で表わせる。したがって、行列

$$R = Z'Z / N = \begin{bmatrix} 1 & r_{12} & \cdots & r_{1n} \\ r_{21} & 1 & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1} & r_{n2} & \cdots & 1 \end{bmatrix} \quad (5)$$

は変量間の相関係数を成分とする相関係数行列となり、 $R$  は対称行列である。なお  $Z'$  は  $Z$  の転置行列を表わす。次に  $n$  個の変量の測定値の 1 次結合としてつくられる新しい変量を

$$f_i = w_1 z_{i1} + w_2 z_{i2} + \cdots + w_n z_{in} = \sum_{j=1}^n w_j z_{ij} \quad (6)$$

と表わし、これを合成変量と呼び、このときの重み  $w_1, w_2, \dots, w_n$  を成分とする列ベクトルを重みベクトルと呼ぶ。

$$w = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} \quad (7)$$

すべてのサンプルに対する合成変量を成分とする列ベクトルを

$$f = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_N \end{bmatrix} = Z w \quad (8)$$

で表わし、この  $f$  を合成変量ベクトルと呼ぶ。合成変量の平均は  $m = 0$ 、分散は  $V = f'f / N = w'Rw$  となる。標準化された合成変量を  $g_i$  で表わすと

$$g_i = f_i / s \quad (9)$$

ここに、 $s$  は合成変量の標準偏差で

$$s = \sqrt{w'Rw} \quad (10)$$

$g$  を標準化された合成変量を成分とする列ベクトルとすると

$$g = Z w / \sqrt{w'Rw} \quad (11)$$

標準化された合成変量を  $g_i$ 、標準化された変量  $j$  の測定値を  $z_{ij}$  とすると、これらの間の相関係数  $a_j$  は

$$a_j = (\sum_{i=1}^N z_{ij} g_i) / N \quad (12)$$

と表わすことができる。1 つの合成変量と  $n$  個の変量との相関係数をまとめたものを

$$a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \quad (13)$$

と表わし、構造ベクトルと呼ぶ。これは

$$a = Z'g / N = R w / \sqrt{w'Rw} \quad (14)$$

となる。 $Z$  から

$$g = Z w_s \quad (15)$$

のように直接  $g$  をつくる重みベクトル  $w_s$  を標準重みベクトルと呼ぶことにする。 $w_s$  を求めるには最も簡単な場合、 $w$  が与えられていれば

$$w_s = w / \sqrt{w'Rw}$$

として得られる。 $w$  が与えられていないで  $a$  が知られている場合は

$$w_s = R^{-1} a \quad (16)$$

として得られる。また、 $w_s$  は標準化された合成変量の分散が 1 であるので

$$w_s'Rw_s = 1 \quad (17)$$

である。

### 3. 主成分分析法

主成分分析法はサンプル行列の中のすべての変量とできるだけ高い相関をもつように合成変量をつくり、同時にサンプル間の差異を示す合成変量の分散を最も大きくするものである。以下に主成分分析法を概説するが、詳細については文献 3) などを参照してほしい。

すべての変量と相関が最も高くなるような解を求めるために

$$V = a'a = a_1^2 + a_2^2 + \cdots + a_n^2 \quad (18)$$

なる  $V$  を求め、これを適当な条件のもとで最大にするように解く。 $a = R w_s$  なる関係を (18) に代入すれば

$$V = w_s'R^2 w_s \quad (19)$$

と書ける。以下表記を簡単にするため  $w_s$  を  $w$  と表わすことにする。 $w'Rw = 1$  という条件のもとで  $V$  を最大にする解を求める。そのためには

$$2T = w'R^2 w - \lambda(w'Rw - 1) \quad (20)$$

なる  $T$  を考え、 $w$  の各成分について偏微分し、その結果をすべて 0 とおく、

$$\partial T / \partial w_j = 0 \quad (j = 1, 2, \dots, n) \quad (21)$$

この結果をまとめて表わすと

$$R^2 w - \lambda R w = 0 \quad (22)$$

となり、 $a = R w$  より

$$R a = \lambda a \quad (23)$$

となる。結局行列  $R$  の固有値問題となり、解の 1 つは次のようにして求められる。②より

$$a' R a = \lambda a' a \quad (24)$$

②の両辺に左から  $w'$  を乗じると

$$w' R^2 w = \lambda w' R w$$

となり、 $w' R w = 1$  より

$$w' R^2 w = \lambda$$

したがって

$$\lambda = V = a' a \quad (25)$$

②4から

$$\lambda = \sqrt{a' R a} \quad (26)$$

が得られる。これと②から

$$a = R a / \sqrt{a' R a} \quad (27)$$

が得られる。この②を用いて  $a$  を逐次的に求めるのが乗べき法といわれるもので、この方法は最大の  $\lambda$  に対する固有ベクトル  $a$  がまずはじめに得られる。②からわかるように解のうちでも最大の  $\lambda$  に対する解をとる必要があるので、乗べき法はそういう点で非常につごうがよい。②は

$$a = R a / \lambda$$

と書け、 $a = R w$  より

$$w = a / \lambda \quad (28)$$

となる。②を②に代入すると

$$R w = \lambda w$$

が得られる。すなわち  $w$  は②の関係を満たす  $R$  の固有ベクトルである。この  $w$  が第 1 合成変量 (第 1 主成分) のための重みベクトルとなる。第 2 以下の合成変量を得るための重みベクトルは、最大の  $\lambda$  に続く固有値に対応する重みベクトルが順次第 2 以下の合成変量の重みベクトルとなる。一般に第  $p$  合成変量 (第  $p$  主成分) は

$$R w_p = \lambda_p w_p \quad (29)$$

を満足する重みベクトルによって与えられる。ここに、 $\lambda_p$  は  $R$  の固有値のうち大きい順に第  $p$  番目の固有値、 $w_p$  はそれに対応した固有ベクトルで、 $w_p$  の大きさは

$$w_p' R w_p = 1 \quad (30)$$

を満たすように定めると、これが標準重みベクトルと

なる。そして

$$a_p = \lambda_p w_p \quad (31)$$

$$\lambda_p = a_p' a_p \quad (32)$$

という関係になる。乗べき法で  $a_p$  を求めるには

$$a_p = R_{p-1} a_p / \sqrt{a_p' R_{p-1} a_p} \quad (33)$$

なる関係から逐次的に求められる。

ここに、

$$R_{p-1} = R - \sum_{i=1}^{p-1} a_i a_i' \quad (34)$$

であり、 $R_{p-1}$  は第  $p-1$  残差行列と呼ばれる。

そして  $\lambda_p$  は

$$\lambda_p = \sqrt{a_p' R_{p-1} a_p} \quad (35)$$

から求め

$$w_p = a_p / \lambda_p \quad (36)$$

によって  $w_p$  を求めればよい。

#### 4. サンプルの次元減少

主成分分析法の最初で述べたように合成変量ベクトル  $g$  の各成分はサンプルの各変量の測定値と最も相関が高いようにつくられており、同時に各成分間の差異を示す分散が最大となるようにつくられている。

$g = Z w$  からわかるように標準重みベクトル  $w$  は各サンプルより特徴量を抽出する特徴抽出ベクトルとみなされ、 $g$  は各サンプルの特徴量を成分とする特徴量ベクトルと考えることができる。さらに行列  $R$  の固有値  $\lambda_i = a_i' a_i$  の大きさは  $\lambda_i$  に対応した特徴抽出ベクトル  $w_i$  の抽出能力および  $w_i$  により抽出された特徴量の有用性についての評価値となっている。

$R$  の固有値が

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m, \quad m = \text{Rank } R$$

のような大きさの順になっているとすれば、これに対応した  $w_i, g_i$  ( $i = 1, 2, \dots, m$ ) もこの順にそれぞれ並べられる。もし  $\lambda_1 \gg \lambda_i$  ( $i = 2, 3, \dots, m$ ) ならば、サンプル行列  $Z$  は  $w_1$  によって抽出されたベクトル  $g_1$  でもって記述され、 $\lambda_1 \geq \lambda_2 \gg \lambda_i$  ( $i = 3, 4, \dots, m$ ) ならば、 $w_1, w_2$  によって抽出される  $g_1, g_2$  でもって  $Z$  を記述できる。すなわち、 $n$  次元のサンプルを 1 個あるいは 2 個の特徴量で記述でき、次元を  $n$  から 1 あるいは 2 に減じることができる。

一般に  $R$  の固有値の大きさが

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_l \gg \lambda_{l+1} \geq \dots$$

のように第  $l+1$  番目の固有値  $\lambda_{l+1}$  から急激に小さくなっているとき、サンプル行列  $Z$  から特徴抽出ベクトル  $w_1, w_2, \dots, w_l$  によって抽出された特徴量ベクトル  $g_1, g_2, \dots, g_l$  でもって  $Z$  は大きさ  $N \times l$

の行列

$$G = [g_1, g_2, \dots, g_t] = Z [w_1, w_2, \dots, w_t] \quad (37)$$

で記述される。

5. クラスタリングアルゴリズム

クラスタリングの評価関数として

$$W = \sum_{i=1}^k S_i \quad (38)$$

ここに、 $k$  はクラスター数で既知

$$S_i = \sum_{Y \in P_i} \|Y - M_i\|^2, \quad M_i = \sum_{Y \in P_i} Y/n_i$$

で、 $P_i$  はクラスター、 $Y$  は  $P_i$  のサンプル、 $n_i$  は  $P_i$  のサンプル数とする。

最初与えられた  $N$  個のサンプルを適当に  $k$  個のクラスターに配属する。クラスタリングを行なうに際して、各クラスターにそれぞれ同じ類のサンプルを集めるには、すなわち各クラスターの  $S_i$  の値ができるだけ最小 (結果として  $W$  が最小) になるようにするためには、サンプルを1つのクラスターから他のどのクラスターに移せばよいかを決定すればよい。もちろんサンプルによってはいまのクラスターが最適のことも当然ありうるであろう。いまクラスター  $P_i$  中のサンプル  $Y_s$  を他のクラスター  $P_j$  に移したとき、平均  $M_i, M_j$  および  $S_i, S_j$  がどのように変化するか求めてみよう。

平均  $M_j$  は

$$M_j^* = M_j + (Y_s - M_j)/(n_j + 1) \quad (39)$$

のように変わり、 $S_j$  は

$$\begin{aligned} S_j^* &= \sum_{Y \in P_j} \|Y - M_j^*\|^2 + \|Y_s - M_j^*\|^2 \\ &= \sum_{Y \in P_j} \|Y - M_j - (Y_s - M_j)/(n_j + 1)\|^2 \\ &\quad + \|(Y_s - M_j)n_j/(n_j + 1)\|^2 \\ &= S_j + \|Y_s - M_j\|^2 n_j/(n_j + 1) \\ &= S_j + \Delta S_j \end{aligned} \quad (40)$$

のように増加する。

平均  $M_i$  は

$$M_i^* = M_i - (Y_s - M_i)/(n_i - 1) \quad (41)$$

のように変わり、 $S_i$  は

$$\begin{aligned} S_i^* &= S_i - \|Y - M_i\|^2 n_i/(n_i - 1) \\ &= S_i - \Delta S_i \end{aligned} \quad (42)$$

のように減少する。

$Y_s$  を  $P_i$  から  $P_j$  に移すことにより評価関数  $W$  の値を減少させるには  $\Delta S_i > \Delta S_j$  でなければならず、 $W$  を最大に減少させるには  $\Delta S_j$  が最小のクラスターを見つければよい。

反復アルゴリズム：

- Step 1.  $N$  個のサンプルを  $k$  個のクラスターに適当に配属し、 $W$  と各クラスターの平均  $M_1, M_2, \dots, M_k$  を計算する。
- Step 2. あるクラスターの中からサンプル  $Y_s$  を選ぶ。
- Step 3.  $\Delta S_i$  および  $\Delta S_j (j \neq i)$  を計算する。
- Step 4.  $\Delta S_i \leq \Delta S_j (j = 1, 2, \dots, k)$  のとき  $Y_s$  を  $P_i$  に移す。
- Step 5.  $W, M_i, M_j$  を更新する。
- Step 6.  $N$  個のサンプルについて Step 2 ~ Step 5 を実行しても  $W$  の値が変化しないなら停止し、そうでなければ Step 2 へ。

6. 実験結果および検討

上述した方法の有用性を確かめるため電子計算機を用いて具体的なパターンを対象にシミュレーションを行なった。以下にその実験結果および検討した結果について述べる。

Table 1 は3次元のサンプルを主成分分析法により、1次元で表わしたものであり、Table 2 は同じく3次元のサンプルを1次元、2次元および3次元で表わしたものである。

Table 1 1-dimensional representation of samples in three dimensions

3-dimensional samples	1-dimensional representation	3-dimensional samples	1-dimensional representation
A <sub>1</sub> (-4, 8, -2)	A <sub>1</sub> (0.62)	B <sub>4</sub> (3, 5, 0)	B <sub>4</sub> (-0.31)
A <sub>2</sub> (-4, 9, -3)	A <sub>2</sub> (0.62)	B <sub>5</sub> (3, 6, -3)	B <sub>5</sub> (-0.45)
A <sub>3</sub> (-4, 10, 1)	A <sub>3</sub> (0.97)	B <sub>6</sub> (3, 7, 5)	B <sub>6</sub> (0.19)
A <sub>4</sub> (-4, 11, 5)	A <sub>4</sub> (1.33)	B <sub>7</sub> (4, 4, 2)	B <sub>7</sub> (-0.36)
A <sub>5</sub> (-5, 7, 1)	A <sub>5</sub> (0.88)	B <sub>8</sub> (4, 5, 1)	B <sub>8</sub> (-0.36)

$A_6(-5, 8, -1)$	$A_6(0.81)$	$B_9(4, 6, -2)$	$B_9(-0.50)$
$A_7(-5, 9, 3)$	$A_7(1.16)$	$B_{10}(4, 7, 0)$	$B_{10}(-0.28)$
$A_8(-5, 10, 0)$	$A_8(1.03)$	$C_1(2, -6, -4)$	$C_1(-1.27)$
$A_9(-5, 11, -1)$	$A_9(1.03)$	$C_2(2, -7, 1)$	$C_2(-1.00)$
$A_{10}(-6, 8, 0)$	$A_{10}(1.00)$	$C_3(2, -8, -3)$	$C_3(-1.35)$
$A_{11}(-6, 9, 4)$	$A_{11}(1.35)$	$C_4(2, -9, -5)$	$C_4(-1.56)$
$A_{12}(-6, 10, 6)$	$A_{12}(1.56)$	$C_5(3, -6, 2)$	$C_5(-0.98)$
$A_{13}(-6, 11, 2)$	$A_{13}(1.36)$	$C_6(3, -7, -2)$	$C_6(-1.33)$
$A_{14}(-7, 9, 0)$	$A_{14}(1.19)$	$C_7(3, -8, 0)$	$C_7(-1.26)$
$A_{15}(-7, 10, -1)$	$A_{15}(1.20)$	$C_8(4, -6, -1)$	$C_8(-1.31)$
$B_1(2, 5, -1)$	$B_1(-0.26)$	$C_9(4, -7, 3)$	$C_9(-1.10)$
$B_2(2, 6, 2)$	$B_2(0.03)$	$C_{10}(4, -8, 4)$	$C_{10}(-1.10)$
$B_3(3, 4, -1)$	$B_3(-0.45)$	$C_{11}(5, -7, 5)$	$C_{11}(-1.08)$

Table 2 1-, 2-, 3-dimensional representation of samples in three dimensions

3-dimensional samples	1-dimensional representation	2-dimensional representation	3-dimensional representation
$A_1(-5, 5, 0)$	$A_1(0.95)$	$A_1(0.95, -0.28)$	$A_1(0.95, -0.28, -0.71)$
$A_2(-5, 6, 1)$	$A_2(1.28)$	$A_2(1.28, 0.31)$	$A_2(1.28, 0.31, -0.37)$
$A_3(-5, 7, 2)$	$A_3(1.61)$	$A_3(1.61, 0.34)$	$A_3(1.61, 0.34, -0.04)$
$A_4(-6, 5, 3)$	$A_4(1.83)$	$A_4(1.83, -0.14)$	$A_4(1.83, -0.14, 0.24)$
$A_5(-6, 6, 0)$	$A_5(1.12)$	$A_5(1.12, 0.34)$	$A_5(1.12, 0.34, -0.86)$
$A_6(-6, 7, -1)$	$A_6(0.94)$	$A_6(0.94, 0.59)$	$A_6(0.94, 0.59, -1.24)$
$A_7(-7, 5, -2)$	$B_7(0.64)$	$A_7(0.64, 0.33)$	$A_7(0.64, 0.33, -1.67)$
$A_8(-7, 6, -3)$	$A_8(0.45)$	$A_8(0.45, 0.58)$	$A_8(0.45, 0.58, -2.05)$
$A_9(-7, 7, 1)$	$A_9(1.55)$	$A_9(1.55, 0.28)$	$A_9(1.55, 0.28, -0.65)$
$A_{10}(-8, 6, 0)$	$A_{10}(1.33)$	$A_{10}(1.33, 0.17)$	$A_{10}(1.33, 0.17, -1.11)$
$B_1(-3, -6, -2)$	$B_1(-0.54)$	$B_1(-0.54, -0.90)$	$B_1(-0.54, -0.90, -0.89)$
$B_2(-3, -7, -1)$	$B_2(-0.35)$	$B_2(-0.35, -1.15)$	$B_2(-0.35, -1.15, -0.51)$
$B_3(-4, -5, 0)$	$B_3(0.15)$	$B_3(0.15, -1.06)$	$B_3(0.15, -1.06, -0.33)$
$B_4(-4, -6, 1)$	$B_4(0.34)$	$B_4(0.34, -1.31)$	$B_4(0.34, -1.31, 0.05)$
$B_5(-4, -7, -4)$	$B_5(-1.02)$	$B_5(-1.02, -0.90)$	$B_5(-1.02, -0.90, -1.71)$
$B_6(-4, -8, -3)$	$B_6(-0.83)$	$B_6(-0.83, -1.15)$	$B_6(-0.83, 1.15, -1.32)$
$B_7(-5, -5, 0)$	$B_7(0.25)$	$B_7(0.25, -1.14)$	$B_7(0.25, -1.14, -0.46)$
$B_8(-5, -6, 1)$	$B_8(0.44)$	$B_8(0.44, -1.39)$	$B_8(0.44, -1.39, -0.08)$
$B_9(-5, -7, 2)$	$B_9(0.63)$	$B_9(0.63, -1.64)$	$B_9(0.63, -1.64, 0.30)$
$B_{10}(-5, -8, 3)$	$B_{10}(0.82)$	$B_{10}(0.82, -1.89)$	$B_{10}(0.82, -1.89, 0.68)$
$C_1(4, 7, 0)$	$C_1(0.18)$	$C_1(0.18, 1.29)$	$C_1(0.18, 1.29, 0.40)$
$C_2(5, 5, 3)$	$C_2(0.72)$	$C_2(0.72, 0.76)$	$C_2(0.72, 0.76, 1.64)$
$C_3(5, 6, 2)$	$C_3(0.53)$	$C_3(0.53, 1.01)$	$C_3(0.53, 1.01, 1.26)$
$C_4(5, 7, 1)$	$C_4(0.34)$	$C_4(0.34, 1.26)$	$C_4(0.34, 1.26, 0.88)$
$C_5(6, 5, -3)$	$C_5(-0.93)$	$C_5(-0.93, 1.50)$	$C_5(-0.93, 1.50, -0.37)$
$C_6(6, 6, -2)$	$C_6(-0.61)$	$C_6(-0.61, 1.53)$	$C_6(-0.61, 1.53, -0.04)$
$C_7(6, 7, -1)$	$C_7(-0.28)$	$C_7(-0.28, 1.56)$	$C_7(-0.28, 1.56, 0.29)$

$C_8(7, 5, 1)$	$C_8(-0.00)$	$C_8(-0.00, 1.14)$	$C_8(-0.00, 1.14, 1.18)$
$C_9(7, 6, 0)$	$C_9(-0.19)$	$C_9(-0.19, 1.39)$	$C_9(-0.19, 1.39, 0.80)$
$C_{10}(7, 7, -4)$	$C_{10}(-1.15)$	$C_{10}(-1.15, 1.98)$	$C_{10}(-1.15, 1.98, -0.65)$
$D_1(4, -6, 0)$	$D_1(-0.73)$	$D_1(-0.73, -0.55)$	$D_1(-0.73, -0.55, 0.72)$
$D_2(4, -7, 1)$	$D_2(-0.54)$	$D_2(-0.54, -0.80)$	$D_2(-0.54, -0.80, 1.10)$
$D_3(5, -5, 2)$	$D_3(-0.24)$	$D_3(-0.24, -0.55)$	$D_3(-0.24, -0.55, 1.53)$
$D_4(5, -6, 3)$	$D_4(-0.05)$	$D_4(-0.05, -0.80)$	$D_4(-0.05, -0.80, 1.91)$
$D_5(5, -7, 4)$	$D_5(0.13)$	$D_5(0.13, -1.05)$	$D_5(0.13, -1.05, 2.29)$
$D_6(6, -5, -1)$	$D_6(-1.12)$	$D_6(-1.12, -0.13)$	$D_6(-1.12, -0.13, 0.59)$
$D_7(6, -6, -2)$	$D_7(-1.45)$	$D_7(-1.45, -0.16)$	$D_7(-1.45, -0.16, 0.26)$
$D_8(6, -7, -3)$	$D_8(-1.77)$	$D_8(-1.77, -0.19)$	$D_8(-1.77, -0.19, -0.07)$
$D_9(7, -6, -4)$	$D_9(-2.06)$	$D_9(-2.06, 0.14)$	$D_9(-2.06, 0.14, -0.33)$
$D_{10}(7, -7, -5)$	$D_{10}(-2.39)$	$D_{10}(-2.39, 0.11)$	$D_{10}(-2.39, 0.11, -0.66)$

Table 3はTable 1のサンプルについてクラスタリングを行なった結果であり、良好な結果が得られている。この例では相関行列  $R$  の固有値は  $\lambda_1=1.74$ ,  $\lambda_2=0.96$  として得られ、 $\lambda_1$  は  $\lambda_2$  にくらべかなり大きい。この事実は  $\lambda_1$  に対応した特徴抽出ベクトル  $w_1$  を用いてサンプル行列  $Z$  から非常に有用な特徴量が抽出できることを意味し、 $\lambda_2$  に対応した  $w_2$  を用いて  $Z$  から抽出される特徴量より、はるかに有用であることを示している。したがって、 $Z$  を  $G=[g_1]=Z[w_1]$  で十分記述でき、1次元で記述されたサンプルに対してクラスタリングを行なっても十分良好な結果が期待できる。固有値の大きさから予想されることと実験結果がよく一致している例といえよう。

Table 4はTable 2のサンプルについてクラスタリングを行なった結果である。表に示されたような1次元に減じたサンプルについてクラスタリングを行なうとエラーが非常に多いことがわかる。ついで2次元に減じたサンプルについて行なってみるとエラーが少し改善されている。そして3次元で行なうと完全にエラーなしでクラスタリングが行なわれている。

これらの実験結果を固有値の大きさから検討してみよう。この例では相関行列  $R$  の固有値が  $\lambda_1=1.19$ ,  $\lambda_2=0.95$ ,  $\lambda_3=0.86$  として得られ、各固有値の間にあまり差がない。このことは  $\lambda_1$ ,  $\lambda_2$  にそれぞれ対応した特徴抽出ベクトル  $w_1$ ,  $w_2$  を用いてサンプル行列  $Z$  から特徴量を抽出し、それによって  $Z$  を記述しても良好なクラスタリングが期待できないことを示して

いる。以上のことから固有値はサンプルを何次元に減じてクラスタリングを行なうのが適当かの目安を与えていることになる。固有値列  $\lambda_1, \lambda_2, \dots$  があるところから急に小さくなるなら、たとえば  $l+1$  番目の固有値から小さくなるなら、サンプルを  $l$  次元に減じてクラスタリングを行なっても良好な結果が期待できると考えられる。

Table 3 Clustering results for samples in Table 1

Initial Partion			Results for 1-dimensional representation		
$P_1$	$P_2$	$P_3$	$P_1$	$P_2$	$P_3$
A <sub>1</sub>	A <sub>2</sub>	A <sub>3</sub>	C <sub>1</sub>	B <sub>1</sub>	A <sub>1</sub> , A <sub>13</sub>
A <sub>4</sub>	A <sub>5</sub>	A <sub>6</sub>	C <sub>2</sub>	B <sub>2</sub>	A <sub>2</sub> , A <sub>14</sub>
A <sub>7</sub>	A <sub>8</sub>	A <sub>9</sub>	C <sub>3</sub>	B <sub>3</sub>	A <sub>3</sub> , A <sub>15</sub>
A <sub>10</sub>	A <sub>11</sub>	A <sub>12</sub>	C <sub>4</sub>	B <sub>4</sub>	A <sub>4</sub>
A <sub>13</sub>	A <sub>14</sub>	A <sub>15</sub>	C <sub>5</sub>	B <sub>5</sub>	A <sub>5</sub>
B <sub>1</sub>	B <sub>2</sub>	B <sub>3</sub>	C <sub>6</sub>	B <sub>6</sub>	A <sub>6</sub>
B <sub>4</sub>	B <sub>5</sub>	B <sub>6</sub>	C <sub>7</sub>	B <sub>7</sub>	A <sub>7</sub>
B <sub>7</sub>	B <sub>8</sub>	B <sub>9</sub>	C <sub>8</sub>	B <sub>8</sub>	A <sub>8</sub>
B <sub>10</sub>	C <sub>1</sub>	C <sub>2</sub>	C <sub>9</sub>	B <sub>9</sub>	A <sub>9</sub>
C <sub>3</sub>	C <sub>4</sub>	C <sub>5</sub>	C <sub>10</sub>	B <sub>10</sub>	A <sub>10</sub>
C <sub>6</sub>	C <sub>7</sub>	C <sub>8</sub>	C <sub>11</sub>		A <sub>11</sub>
C <sub>9</sub>	C <sub>10</sub>	C <sub>11</sub>			A <sub>12</sub>

Table 4 Clustering results for 1-, 2-, 3-dimensional representation of samples in Table 2

Initial partition				Results for 1-dimensional representation				Results for 2-dimensional representation				Results for 3-dimensional representation			
P <sub>1</sub>	P <sub>2</sub>	P <sub>3</sub>	P <sub>4</sub>	P <sub>1</sub>	P <sub>2</sub>	P <sub>3</sub>	P <sub>4</sub>	P <sub>1</sub>	P <sub>2</sub>	P <sub>3</sub>	P <sub>4</sub>	P <sub>1</sub>	P <sub>2</sub>	P <sub>3</sub>	P <sub>4</sub>
A <sub>1</sub>	A <sub>2</sub>	A <sub>3</sub>	A <sub>4</sub>	D <sub>7</sub>	A <sub>1</sub>	C <sub>5</sub>	B <sub>8</sub>	C <sub>1</sub>	A <sub>1</sub>	D <sub>1</sub>	B <sub>1</sub>	C <sub>1</sub>	A <sub>1</sub>	D <sub>1</sub>	B <sub>1</sub>
A <sub>5</sub>	A <sub>6</sub>	A <sub>7</sub>	A <sub>8</sub>	D <sub>8</sub>	A <sub>2</sub>	C <sub>6</sub>	B <sub>4</sub>	C <sub>4</sub>	A <sub>2</sub>	D <sub>6</sub>	B <sub>2</sub>	C <sub>2</sub>	A <sub>2</sub>	D <sub>2</sub>	B <sub>2</sub>
A <sub>9</sub>	A <sub>10</sub>	B <sub>1</sub>	B <sub>2</sub>	D <sub>9</sub>	A <sub>3</sub>	C <sub>7</sub>	B <sub>7</sub>	C <sub>5</sub>	A <sub>3</sub>	D <sub>7</sub>	B <sub>3</sub>	C <sub>3</sub>	A <sub>3</sub>	D <sub>3</sub>	B <sub>3</sub>
B <sub>3</sub>	B <sub>4</sub>	B <sub>5</sub>	B <sub>6</sub>	D <sub>10</sub>	A <sub>4</sub>	C <sub>9</sub>	B <sub>8</sub>	C <sub>6</sub>	A <sub>4</sub>	D <sub>8</sub>	B <sub>4</sub>	C <sub>4</sub>	A <sub>4</sub>	D <sub>4</sub>	B <sub>4</sub>
B <sub>7</sub>	B <sub>8</sub>	B <sub>9</sub>	B <sub>10</sub>		A <sub>5</sub>	C <sub>10</sub>	B <sub>9</sub>	C <sub>7</sub>	A <sub>5</sub>	D <sub>9</sub>	B <sub>6</sub>	C <sub>5</sub>	A <sub>5</sub>	D <sub>5</sub>	B <sub>5</sub>
C <sub>1</sub>	C <sub>2</sub>	C <sub>3</sub>	C <sub>4</sub>		A <sub>6</sub>	B <sub>1</sub> *	B <sub>10</sub>	C <sub>8</sub>	A <sub>6</sub>	D <sub>10</sub>	B <sub>7</sub>	C <sub>6</sub>	A <sub>6</sub>	D <sub>6</sub>	B <sub>6</sub>
C <sub>5</sub>	C <sub>6</sub>	C <sub>7</sub>	C <sub>8</sub>		A <sub>9</sub>	B <sub>2</sub> *	A <sub>7</sub> *	C <sub>9</sub>	A <sub>7</sub>	B <sub>5</sub> *	B <sub>8</sub>	C <sub>7</sub>	A <sub>7</sub>	D <sub>7</sub>	B <sub>7</sub>
C <sub>9</sub>	C <sub>10</sub>	D <sub>1</sub>	D <sub>2</sub>		A <sub>10</sub>	B <sub>5</sub> *	A <sub>8</sub> *	C <sub>10</sub>	A <sub>8</sub>		B <sub>9</sub>	C <sub>8</sub>	A <sub>8</sub>	D <sub>8</sub>	B <sub>8</sub>
D <sub>3</sub>	D <sub>4</sub>	D <sub>5</sub>	D <sub>6</sub>			B <sub>6</sub> *	C <sub>1</sub> *		A <sub>9</sub>		B <sub>10</sub>	C <sub>9</sub>	A <sub>9</sub>	D <sub>9</sub>	B <sub>9</sub>
D <sub>7</sub>	D <sub>8</sub>	D <sub>9</sub>	D <sub>10</sub>			D <sub>1</sub> *	C <sub>2</sub> *		A <sub>10</sub>		D <sub>2</sub> *	C <sub>9</sub>	A <sub>10</sub>	D <sub>10</sub>	B <sub>10</sub>
						D <sub>2</sub> *	C <sub>3</sub> *		C <sub>2</sub> *		D <sub>3</sub> *	C <sub>10</sub>			
						D <sub>3</sub> *	C <sub>4</sub> *		C <sub>3</sub> *		D <sub>4</sub> *				
						D <sub>6</sub> *	C <sub>5</sub> *				D <sub>5</sub> *				
						D <sub>4</sub> *									
						D <sub>5</sub> *									

## 7. ま と め

以上主成分分析法を用いてサンプルを低次元に減じ、それに対しクラスターリングを行なう教師なし反復アルゴリズムについて述べた。そして相関行列  $R$  の固有値の大きさからサンプルを何次元に減じてクラスターリングを行なえるか予想できることがわかった。

本論文で示した電子計算機を用いた実験では簡単な3次元の人工的なサンプルであるが、実際のパターン類別においてはサンプルの次元が非常に大きく(100次元以上になることも珍しくない)、本方法がこのような場合にも有効に働くことが確かめられれば非常に有用なものとなろう。当然サンプルのための記憶領域、クラスターリングに要する時間は飛躍的に節減されることになる。今後本方法の有用性を確かめるべく種々の高次元サンプルについて実験を行なっていく

い。

## 謝 辞

本研究を進めるにあたり、種々の討論や意見を寄せて下さった山口大学工学部電子工学科助手金岡泰保氏に深謝する。なお、本研究には渡辺記念学術奨励会の奨学金の補助を受けて行なったことを付記する。

## 参 考 文 献

- 1) R.O Duda, P.E. Hart: "PATTERN CLASSIFICATION AND SCENE ANALYSIS" John Wiley & Sons, New York (1973)
- 2) K.S. Fu: "Digital Pattern Recognition" Springer, Berlin (1976)
- 3) 芝祐順: "行動科学における相関分析法" 東大出版会 (1975)

(昭和52年10月10日受理)