

## 因子分析法による特徴抽出†

岡田 敏彦†† 富田 真吾††

本論文は因子分析法における単純構造の概念を基にし、パターンの特徴抽出問題を論じている。すなわち  $n$  次元空間上のパターンクラスは、クラス間の相違が各クラスのパターンの  $n$  個の変量の観測値に依存するが、常に  $n$  個の変量のすべてが同程度にクラス間の相違に影響を及ぼしているとはかぎらない。このため、各クラスにおいては他のクラスとの相違に大きな役割を果たしている変量グループとそうではない変量グループが存在する。この事実に着目し、クラス間の相違に大きな役割を果たしている変量グループを明らかにする目的で因子分析法の単純構造化の手法を用いることにより、これらの変量グループにもっとも相関の強い特徴を抽出する単純構造系を提案している。単純構造系の座標軸はそれぞれ各クラスに対応し、クラスの類別に重要な情報を担い、クラスが  $K$  個分布しているなら、高々  $K$  個の座標軸よりなる系として構成される。さらに具体的なパターンをもちいて計算機シミュレーションを行い、従来の Karhunen-Loève 系との比較検討し、本方法の有用性について述べている。

## 1. ま え が き

パターンの特徴抽出法として従来より種々の方法が提案されてきている。パターンの特徴抽出の目的はパターンの膨大な観測値、たとえば一つの波形や画像をデジタル化するとしばしば膨大な量の観測値集合となるが、これらのものからパターンを識別する情報を十分含むより小さな特徴量集合に、もとのパターン観測値を写像する変換を見つけることである。よく知られた方法の一つに Karhunen-Loève 変換 (K-L 変換) による方法、すなわち、正規直交固有ベクトル変換<sup>1,2)</sup>がある。この方法はパターン自己相関行列の固有ベクトルに対応する固有値の大きい順に選ばれ、選ばれた固有ベクトルがつくる部分空間にもとのパターンを射影するものである。この部分空間には各パターンにもっとも相関の強い特徴量が抽出され、この点に関しては最適な空間であるが、パターンを類別するための特徴抽出空間としては必ずしも最適ではない。この欠点を補うため 1970 年 Fukunaga-Koontz<sup>3)</sup> は修正された K-L 変換法を提案した。すなわち 2 クラス類別問題において二つのクラスの重みつきパターン自己相関行列をそれぞれ  $R_1, R_2$  とし、これらの和  $R_0 = R_1 + R_2$  に対してある予備変換  $PR_0P^T = I$  ( $P$  は変換行列,  $I$  は単位行列,  $T$  は転置を示す) を行うとき、変換された両クラスの自己相関行列  $PR_1P^T, PR_2P^T$  は同じ固有ベクトルと固有値をもつ。クラス

1 に対する固有値が  $1 \geq \lambda_1^{(1)} \geq \lambda_2^{(2)} \geq \dots \geq \lambda_n^{(1)} \geq 0$  のように並べられるとすれば、クラス 2 に対する固有値は  $\lambda_i^{(2)} = 1 - \lambda_i^{(1)}$  となる。したがって  $\lambda_i^{(1)}$  に対応する固有ベクトルはクラス 1 のもっとも重要な特徴を抽出するがクラス 2 に対してはもっとも重要でない特徴を抽出する。逆に  $\lambda_n^{(2)} (= 1 - \lambda_n^{(1)})$  に対応する固有ベクトルはクラス 2 のもっとも重要な特徴を抽出するがクラス 1 に対してはもっとも重要でない特徴を抽出する。すなわち両クラスは共通に重要な特徴は共有できなくなる。Fukunaga らは  $|\lambda_i^{(1)} - 0.5|$  の値の大きい順に対応する固有ベクトルを選び、選ばれた固有ベクトルがつくる部分空間にパターンを射影することを提案している。これは非常にすぐれた方法であるが一般の多クラス問題に拡張できないのが欠点である。1975 年 Foley-Sammon<sup>4)</sup> は 2 クラス問題に対し二つのクラス間の距離をできるだけ大きく、同時にクラス内のパターンの分散をできるだけ小さくする変換ベクトルを見つけ、これらのベクトルがつくる部分空間にパターンを射影する方法を提案している。そして Fukunaga-Koontz 変換法ではうまく類別できない例についても良好に類別することを示している。しかしやはり 2 クラス問題に限定され一般性に欠けている。1965 年 S. Watanabe<sup>1)</sup> は K-L 系のパターン認識への応用にもとづき一般的な因子分析法による特徴抽出への応用について示唆を与えている。本論文ではこの示唆をもとに 3 個以上の多クラス問題について考察した。すなわち  $n$  次元空間に  $K$  個のパターンクラス  $C_i$  ( $i=1, \dots, K$ ) が分布しているとしよう。各クラスはそれぞれ特有の特徴を有していて、これらの特徴は  $n$  個の変量のうち主として一部の変量グループによって観測され

† Feature Extraction of Patterns by Factor Analysis by TOSHIHIKO OKADA and SHINGO TOMITA (Faculty of Engineering, Yamaguchi University).

†† 山口大学工学部電子工学科

ているものとする. すなわちあるクラス  $C_i$  の特徴は主として  $n_i$  個の変量が, またあるクラス  $C_j$  の特徴は別の  $n_j$  個の変量が主に観測している場合である (付録 [1] を参照). しかしこれらの変量グループは事前にはわかっていないとしよう. このようなとき因子分析法における単純構造化の概念を導入し, 変量のグループ分けを行い, 各変量グループにもっとも相関の強い方向ベクトルを見つけだし, それらのベクトルがつくる部分空間にパターンを射影するならばその空間は  $K$  個のパターンクラスの類別に有用な特徴を抽出するであろう. しかしここで問題は因子分析法ではパターンを各変量について平均が 0, 分散が 1 となるように基準化することである. 平均を 0 にする操作は単に空間の原点の移動するだけであり, パターンの分布形状になんの影響も与えないが, 分散を 1 にする操作はパターンの分布形状を変えてしまう. これはパターンの特徴抽出を考えるうえで非常に問題になる. したがって因子分析法の理論をそのまま導入するには問題がある. そこで本論文では理論を展開するうえで式を簡潔に表現するためと, パターンの分布形状に影響を与えないとの理由から各変量について平均が 0 になるように前処理するが, 分散についてはそのままにする. そのうえで因子分析法の単純構造化の理論を特徴抽出への応用のために再構成する.

2章で数学的準備と特徴抽出機構の基本原則を, 3章で特徴抽出のための単純構造化変換理論を述べる. そして4章で K-L 変換との相違について述べ, 5章で本変換法の有用性を確かめるため計算機実験の結果を示し, 同じデータを使って行った K-L 変換の結果と比較検討をする.

## 2. 準備

$n$  次元空間に  $N$  個のパターン集合  $\{x_1, x_2, \dots, x_N\}$  が与えられ, 原点は  $\sum_{i=1}^N x_i = 0$  となるようにとられているものとしよう. すなわち

$$\sum_{i=1}^N x_{ij} = 0 \quad (j=1, \dots, n) \quad (1)$$

である. ここに,  $x_i = (x_{i1}, x_{i2}, \dots, x_{in})^T$ ,  $x_{ij}$  はパターン  $x_i$  の変量  $j$  の観測値を表わし実数である.

パターン行列を

$$X = (x_1, x_2, \dots, x_N)^T, \quad (2)$$

パターン自己相関行列を

$$R = E[x_i x_i^T] = X^T X / N \quad (3)$$

として表わす. パターン行列  $X$  よりベクトル  $\xi_1$  によって抽出される特徴量を成分とするベクトルを

$$p_1 = X \xi_1 \quad (4)$$

と表わす.  $p_1$  の成分  $p_{1i} (i=1, \dots, N)$  はパターン  $x_i$  より抽出された特徴量を示す.  $p_1$  の成分の平均は 0, 分散は  $\xi_1^T R \xi_1$  となる.  $p_1$  と各変量との相関係数を成分とするベクトル  $\eta_1$  は,  $p_1$  を基準化したものを

$$q_1 = p_1 / \sqrt{\xi_1^T R \xi_1} \quad (5)$$

と表わし,  $X$  の各列について, すなわち各変量について基準化したものを

$$X_i = X S^{-1} \quad (6)$$

と表わせば

$$\begin{aligned} \eta_1 &= X_i^T q_1 / N \\ &= S^{-1} R \xi_1 / \sqrt{\xi_1^T R \xi_1} \end{aligned} \quad (7)$$

となる. ここに行列  $S$  は各変量の標準偏差を対角成分とするもので  $S = \text{diag}(s_1, s_2, \dots, s_n)$  であり  $s_i$  は変量  $i$  の標準偏差である. なお  $s_i \neq 0 (i=1, \dots, n)$  とする.  $\eta_1$  は特徴量の構造を表わし, 因子分析法における因子負荷ベクトルに相当するが, ここでは構造ベクトルとよぶことにする.  $\eta_1$  の  $i$  番目の成分  $\eta_{1i}$  は  $p_1$  と変量  $i$  との相関係数を示す.

$p_1$  はパターン間のある種の特徴差を, ひいてはクラス間の特徴差を表わしうるが, さらに  $p_1$  では表わしえなかった別種の特徴差を表わしたいことがある. それを  $p_2$  で表わしたいとき,  $p_2$  は  $p_1$  とはまったく相関のない独立したものにすることがある. すなわち  $p_2$  を  $p_1$  と直交化して

$$p_1^T p_2 = 0$$

となるようにしなければならない. このような  $p_2$  は

$$p_2 = X(I - \xi_1 \eta_1^T S / \sqrt{\xi_1^T R \xi_1}) \xi_2^* \quad (8)$$

$$= X \xi_2^* \quad (9)$$

$$= X \xi_2 \quad (10)$$

として得られる (証明は付録 [2]).

ここに,

$$X_1 = X(I - \xi_1 \eta_1^T S / \sqrt{\xi_1^T R \xi_1}) \quad (11)$$

$$\xi_2 = (I - \xi_1 \eta_1^T S / \sqrt{\xi_1^T R \xi_1}) \xi_2^* \quad (12)$$

で,  $\xi_2^*$  は任意ベクトルである.

この  $X_1$  をもちいて

$$R_1 = X_1^T X_1 / N = R - S \eta_1 \eta_1^T S \quad (13)$$

とおく.  $p_2$  の成分の平均は 0, 分散は  $\xi_2^{*T} R_1 \xi_2^*$  となるので  $p_2$  と各変量との相関係数を成分とするベクトル  $\eta_2$  は,  $p_2$  を基準化して

$$q_2 = p_2 / \sqrt{\xi_2^{*T} R_1 \xi_2^*} \quad (14)$$

とおけば

$$\begin{aligned} \eta_2 &= \mathbf{X}_1^T \mathbf{q}_2 / N \\ &= \mathbf{S}^{-1} \mathbf{R}_1 \xi_2^* / \sqrt{\xi_2^{*T} \mathbf{R}_1 \xi_2^*} \\ &= \mathbf{S}^{-1} \mathbf{R}_2 \xi_2 / \sqrt{\xi_2^T \mathbf{R}_2 \xi_2} \end{aligned} \quad (15)$$

と表わされる。η<sub>2</sub> の *i* 番目の成分 η<sub>2*i*</sub> は **p**<sub>2</sub> と変数 *i* との相関係数を示す。

一般に **p**<sub>1</sub>, **p**<sub>2</sub>, ..., **p**<sub>*i*-1</sub> と直交する **p**<sub>*i*</sub> は

$$\mathbf{p}_i = \mathbf{X} \left( \mathbf{I} - \sum_{j=1}^{i-1} \xi_j \eta_j^T \mathbf{S} / \sqrt{\xi_j^T \mathbf{R}_j \xi_j} \right) \xi_i^* \quad (17)$$

$$= \mathbf{X}_{i-1} \xi_i^* \quad (18)$$

$$= \mathbf{X} \xi_i \quad (19)$$

となる (証明は付録 [3])。ここに、

$$\mathbf{X}_{i-1} = \mathbf{X} \left( \mathbf{I} - \sum_{j=1}^{i-1} \xi_j \eta_j^T \mathbf{S} / \sqrt{\xi_j^T \mathbf{R}_j \xi_j} \right) \quad (20)$$

$$\xi_i = \left( \mathbf{I} - \sum_{j=1}^{i-1} \xi_j \eta_j^T \mathbf{S} / \sqrt{\xi_j^T \mathbf{R}_j \xi_j} \right) \xi_i^* \quad (21)$$

で、ξ<sub>*i*</sub><sup>\*</sup> は任意ベクトルである。また η<sub>*i*</sub> は

$$\eta_i = \mathbf{S}^{-1} \mathbf{R}_{i-1} \xi_i^* / \sqrt{\xi_i^{*T} \mathbf{R}_{i-1} \xi_i^*} \quad (22)$$

$$= \mathbf{S}^{-1} \mathbf{R}_i \xi_i / \sqrt{\xi_i^T \mathbf{R}_i \xi_i} \quad (23)$$

ここに、

$$\begin{aligned} \mathbf{R}_{i-1} &= \mathbf{X}_{i-1}^T \mathbf{X}_{i-1} / N \\ &= \mathbf{R} - \mathbf{S} \sum_{j=1}^{i-1} \eta_j \eta_j^T \mathbf{S} \end{aligned} \quad (24)$$

である。ところで(20)式は

$$\mathbf{X}_{i-1} = \mathbf{X} - \sum_{j=1}^{i-1} \mathbf{p}_j \eta_j^T \mathbf{S} / \sqrt{\xi_j^T \mathbf{R}_j \xi_j}$$

と書くことができ、さらに

$$\mathbf{r}_j^T = \eta_j^T \mathbf{S} / \sqrt{\xi_j^T \mathbf{R}_j \xi_j}$$

とおけば

$$\mathbf{X}_{i-1} = \mathbf{X} - \sum_{j=1}^{i-1} \mathbf{p}_j \mathbf{r}_j^T$$

と書ける。したがって、**X** は

$$\mathbf{X} = \mathbf{p}_1 \mathbf{r}_1^T + \mathbf{p}_2 \mathbf{r}_2^T + \dots + \mathbf{p}_m \mathbf{r}_m^T + \dots$$

と表わされる。**X** の階数を *m* とすれば **X** は *m* 個の互いに直交するベクトルで表わすことができ<sup>6)</sup>、いま **p**<sub>*i*</sub> と **p**<sub>*j*</sub> (*i* ≠ *j*) は互いに直交するベクトルであるので、結局 **X** は

$$\mathbf{X} = \mathbf{p}_1 \mathbf{r}_1^T + \mathbf{p}_2 \mathbf{r}_2^T + \dots + \mathbf{p}_m \mathbf{r}_m^T \quad (25)$$

と表わされる。**X**<sup>T</sup> の階数も *m* であるから **R** = **X**<sup>T</sup>**X**/*N* の階数は *m* となる。したがって **R** は

$$\begin{aligned} \mathbf{R} &= \mathbf{X}^T \mathbf{X} / N \\ &= \mathbf{S} \sum_{i=1}^m \eta_i \eta_i^T \mathbf{S} \end{aligned} \quad (26)$$

と表わされる。

一方特徴抽出ベクトル ξ<sub>*i*</sub> は方向のみが重要であるから正規化したものを

$$\phi_i = \xi_i / \|\xi_i\| \quad (\|\phi_i\|^2 = \phi_i^T \phi_i = 1) \quad (27)$$

とおく。さらに φ<sub>*i*</sub> に対応した η<sub>*i*</sub> をもちいて

$$\gamma_i = \eta_i^T \phi_i \quad (28)$$

を定義する。η<sub>*i*</sub> の成分は抽出された特徴量と各変数との相関係数であり、その二乗和を γ<sub>*i*</sub> は示すので、いま γ<sub>*i*</sub> が大きいことは γ<sub>*i*</sub> に対応した φ<sub>*i*</sub> によって抽出される特徴量が変数と強い相関をもつことになる。すべての変数と、あるいは一部の変数と強い相関をもつかは φ<sub>*i*</sub> の設計にかかってくる。なぜなら φ<sub>*i*</sub> は(21)式からわかるように任意ベクトル ξ<sub>*i*</sub><sup>\*</sup> に依存するからである。したがって、γ<sub>*i*</sub> の大きい順に対応する φ<sub>*i*</sub> を選び、それらがつくる部分空間にパターンを射影するならば、すべての変数、あるいは一部の變数と相関の強い特徴が抽出されることになる。選ばれた φ<sub>*i*</sub> を φ<sub>1</sub>, φ<sub>2</sub>, ..., φ<sub>*i*</sub> とすれば

$$\Phi = (\phi_1, \phi_2, \dots, \phi_i)^T \quad (29)$$

は

$$\mathbf{y} = (y_1, y_2, \dots, y_i)^T = \Phi \mathbf{x} \quad (30)$$

のような変換を行うことになり

$$y_i = \phi_i^T \mathbf{x} \quad (31)$$

はパターン **x** の新しい空間における *i* 番目の特徴量となる。この特徴抽出機構を図1に示す。

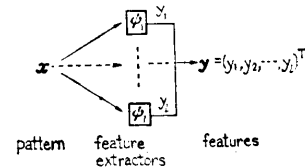


図1 特徴抽出機構

Fig. 1 Feature extraction mechanism.

また φ<sub>*i*</sub> は任意ベクトル ξ<sub>*i*</sub><sup>\*</sup> に依存していることから、φ<sub>*i*</sub>(ξ<sub>*i*</sub><sup>\*</sup>) を可変することにより特徴量構造を制御することが可能となる。ここではパターンの類別のための特徴量構造にするにはどのように φ<sub>*i*</sub>(ξ<sub>*i*</sub><sup>\*</sup>) を設計すればよいかについて考え、次の3章で詳しく述べる。

図2は φ<sub>*i*</sub>(ξ<sub>*i*</sub><sup>\*</sup>) による y<sub>*i*</sub> の制御を図示したものである。

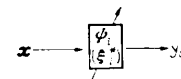


図2 可変 φ<sub>*i*</sub>(ξ<sub>*i*</sub><sup>\*</sup>) による y<sub>*i*</sub> の制御

Fig. 2 Control of y<sub>*i*</sub> by a variable φ<sub>*i*</sub>(ξ<sub>*i*</sub><sup>\*</sup>).

### 3. 単純構造変換

$n$ 次元空間に  $K$  個のパターンクラス  $C_i$  ( $i=1, \dots, K$ ) が分布し、各クラスの特徴は主に一部の変量グループによって観測されているものとしよう (付録 [1] 参照). すなわち、あるクラス  $C_i$  の特徴は  $n$  個の変量のうち主に  $n_i$  個が、またあるクラス  $C_j$  の特徴は別の  $n_j$  個が観測しているといった場合である. このとき各変量グループにもっとも相関の強い特徴量を抽出する変換ベクトルを見つけたし、それらのベクトルがつくる部分空間にパターンを射影するならばクラスを類別する重要な特徴が抽出されるであろう. しかし変量グループがクラスの特徴を主に観測しているとわかっているものの、どのようなグループになっているか事前には判明していないとしよう. このようなとき因子分析法における単純構造化の概念を導入し [定義 1] で示す基準量  $V$  を定め、これを最大にすることによって変量グループを明らかにする.

[定義 1]  $\eta$  の成分の平方  $\eta_j^2$  ( $j=1, \dots, n$ ) の分散

$$V = E[(\eta_j^2 - \mu)^2] = E[\eta_j^4] - \mu^2 \quad (32)$$

$$\text{ここに, } \mu = E[\eta_j^4]$$

を単純構造基準とよび、 $V$  の値を最大にする特徴量の構造を単純構造とよぶ.

$V$  が最大にされると  $\eta$  の成分の絶対値  $|\eta_j|$  が 1 に近いものと 0 に近いものとに分かれ、0.5 程度の値はなくなってしまふ. つまり特徴量の構造は一部の変量グループととくに相関が強くなり、他の変量グループとはとくに相関が弱くなり、中程度の相関がなくなってしまう. そして  $|\eta_j|$  の値が 1 に近い変量がグループとしてとりだされる. 定義より次の補題, 定理を得る.

[補題 1] パターン行列  $X$  より単純構造をなす特徴量ベクトル  $p_1 = X\xi_1$  を抽出するベクトル  $\xi_1$  は

$$\xi_1 = S^{-1}\zeta_1 \quad (33)$$

ここに,

$$\zeta_1 = \eta_1^{(3)} - \eta_1^T \eta_1 \eta_1 / n \quad (34)$$

$$\eta_1 = S^{-1} R S^{-1} \zeta_1 / \sqrt{\zeta_1^T S^{-1} R S^{-1} \zeta_1} \quad (35)$$

$$\eta_1 = (\eta_{11}, \eta_{21}, \dots, \eta_{n1})^T$$

$$\eta_1^{(3)} = (\eta_{11}^3, \eta_{21}^3, \dots, \eta_{n1}^3)^T$$

となる (証明は付録 [4]).

[補題 2]  $p_1$  と直交し、かつ単純構造をなす  $p_2 = X\xi_2$  を抽出するベクトル  $\xi_2$  は

$$\xi_2 = (I - \xi_1 \eta_1^T S / \sqrt{\xi_1^T R \xi_1}) S^{-1} \zeta_2 \quad (36)$$

ここに,

$$\zeta_2 = \eta_2^{(3)} - \eta_2^T \eta_2 \eta_2 / n \quad (37)$$

$$\eta_2 = S^{-1} R_1 S^{-1} \zeta_2 / \sqrt{\zeta_2^T S^{-1} R_1 S^{-1} \zeta_2} \quad (38)$$

$$\eta_2 = (\eta_{12}, \eta_{22}, \dots, \eta_{n2})^T$$

$$\eta_2^{(3)} = (\eta_{12}^3, \eta_{22}^3, \dots, \eta_{n2}^3)^T$$

となる (証明は付録 [4], [5] を参照).

一般に次の定理が成立する.

[定理 1]  $p_1, p_2, \dots, p_{i-1}$  と直交し、かつ単純構造をなす  $p_i = X\xi_i$  を抽出するベクトル  $\xi_i$  は

$$\xi_i = \left( I - \sum_{j=1}^{i-1} \xi_j \eta_j^T S / \sqrt{\xi_j^T R \xi_j} \right) S^{-1} \zeta_i \quad (39)$$

ここに,

$$\zeta_i = \eta_i^{(3)} - \eta_i^T \eta_i \eta_i / n \quad (40)$$

$$\eta_i = S^{-1} R_{i-1} S^{-1} \zeta_i / \sqrt{\zeta_i^T S^{-1} R_{i-1} S^{-1} \zeta_i} \quad (41)$$

$$\eta_i = (\eta_{1i}, \eta_{2i}, \dots, \eta_{ni})^T$$

$$\eta_i^{(3)} = (\eta_{1i}^3, \eta_{2i}^3, \dots, \eta_{ni}^3)^T$$

となる (証明は付録 [4], [5] を参照).

各  $\xi_i$  方向の軸は各クラスのもっとも重要な情報を担うことになる.  $\xi_i$  を正規化したものを  $\phi_i$  とし、 $\gamma$  値の大きい順に選び、選ばれた  $\phi_i$  がつくる部分空間にパターンを射影するのが単純構造変換である. 選ばれた  $l$  個の  $\phi$  を  $\phi_1, \phi_2, \dots, \phi_l$  とすれば

$$\Phi = (\phi_1, \phi_2, \dots, \phi_l)^T \quad (42)$$

はパターン  $x$  を

$$y = \Phi x \quad (43)$$

のように変換する. この  $\phi_i$  を単純化ベクトルとよぶことにする. 一方  $\xi_i$  を求めるには (39) 式からわかるように  $\zeta_i$  を求めねばならない.  $\zeta_i$  は (40) 式から求められるが右辺がまた  $\zeta_i$  の関数となっているので、最初  $\zeta_i$  は適当な初期値を代入し、新しく得られた  $\zeta_i$  を再び右辺の  $\zeta_i$  に代入するといった逐次反復計算を行い所要の精度まで収束させる必要がある. すなわち反復公式

$${}_{p+1}\zeta_i = {}_p\eta_i^{(3)} - {}_p\eta_i^T {}_p\eta_i {}_p\eta_i / n \quad (44)$$

$${}_p\eta_i = S^{-1} R_{i-1} S^{-1} {}_p\zeta_i / \sqrt{{}_p\zeta_i^T S^{-1} R_{i-1} S^{-1} {}_p\zeta_i} \quad (45)$$

をもちいて初期値  ${}_0\zeta_i$  から出発し  ${}_1\zeta_i, {}_2\zeta_i, \dots$  を計算し,

$$\max_j |{}_{p+1}\zeta_{ji} - {}_p\zeta_{ji}| < \epsilon$$

を満足するまで反復計算すればよい. なお  ${}_p\zeta_i = ({}_p\zeta_{1i}, {}_p\zeta_{2i}, \dots, {}_p\zeta_{ni})^T$  である.

### 4. K-L 変換との相違

3章で述べた単純構造変換と K-L 変換との相違に

ついて述べておく。K-L 変換は  $K$  個のパターンクラス  $C_i (i=1, \dots, K)$  が存在し、 $C_i$  の生起確率  $P_i$  を  $C_i$  のパターン自己相関行列に乗じた行列を

$$R^{(i)} = P_i E[\mathbf{x}_j^{(i)} \mathbf{x}_j^{(i)T}] \quad (i=1, \dots, K)$$

とすると、これらの行列を混合したものを

$$R = \sum_{i=1}^K R^{(i)}$$

とおき、この  $R$  の固有ベクトルがもちいられる。 $R$  は対称行列なので固有ベクトルどうしは互いに直交し、固有値は正または 0 である。固有値の大きい順に対応する固有ベクトルが選ばれ、選ばれた固有ベクトルがつくる部分空間にパターンを射影するのが K-L 変換である。各固有ベクトル方向の軸はすべての変量ともっとも相関の強い情報を、すなわち全パターンクラスともっとも相関の強い情報を担うことになる。K-L 変換は全パターンクラスにもっとも相関の強い空間を形成するといえよう。一方単純構造変換は各パターンクラスの特徴がそれぞれ別個の変量グループによって観測されているならば各パターンクラスにもっとも相関の強い特徴空間を形成した。したがって全パターンクラスの特徴が別個の変量グループによって観測されているときは K-L 変換よりも単純構造変換の方がクラス類別のための特徴をよく抽出する。

5. 計算機実験および検討

単純構造変換の有用性を確かめるため人工的なパターンをもちいて計算機実験を行った。表 1 は 4 次元の空間に 39 個のパターンが三つのクラスをなして分布している場合である。クラス  $C_1$  と  $C_2$  の特徴は主に変量 1 と変量 2 が、 $C_3$  は変量 3 と変量 4 が観測している。表 2 の構造ベクトル  $\eta_1$  と  $\eta_2$  は特徴量の構造が単純化されるのをよく示している。さらに表 3 の単純化ベクトル  $\phi_1$  はクラス  $C_1$  と  $C_2$  の、 $\phi_2$  は  $C_3$  の重要な特徴を抽出することをよく示している。表 4 は

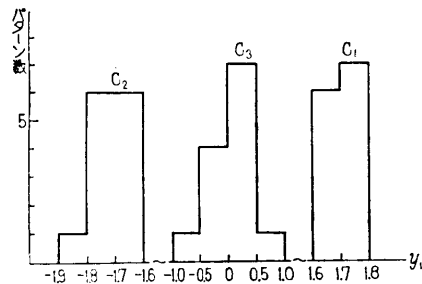


図 3 単純構造変換による特徴量  $y_1$  のヒストグラム  
Fig. 3 Histogram for feature  $y_1$  by simple-structure transformation.

表 1 4 次元パターン  
Table 1 4-dimensional patterns.

Class		$X_1$	$X_2$	$X_3$	$X_4$
$C_1$	1	1.227	1.211	-0.650	-0.666
	2	1.203	1.218	-1.050	-1.066
	3	1.155	1.168	-0.692	-0.674
	4	1.179	1.187	-0.251	-0.264
	5	1.203	1.206	-1.468	-1.432
	6	1.250	1.226	-1.050	-1.051
	7	1.203	1.233	-0.657	-0.645
	8	1.215	1.247	-0.271	-0.293
	9	1.234	1.238	0.097	0.110
	10	1.177	1.235	-1.040	-1.019
	11	1.186	1.206	-0.664	-0.670
	12	1.200	1.171	-0.299	-0.282
	13	1.203	1.159	-0.657	-0.670
$C_2$	14	-1.223	-1.122	-0.664	-0.681
	15	-1.211	-1.162	-1.040	-1.069
	16	-1.215	-1.181	-0.671	-0.670
	17	-1.199	-1.169	-0.320	-0.300
	18	-1.175	-1.189	-1.433	-1.396
	19	-1.187	-1.201	-1.016	-1.001
	20	-1.246	-1.210	-0.661	-0.672
	21	-1.251	-1.153	-0.282	-0.289
	22	-1.282	-1.165	0.080	0.046
	23	-1.270	-1.201	-1.019	-1.051
	24	-1.230	-1.210	-0.664	-0.659
	25	-1.175	-1.409	-0.317	-0.275
	26	-1.170	-1.363	-0.657	-0.681
	$C_3$	27	0.021	0.008	1.279
28		-0.231	-0.248	1.422	1.121
29		0.009	-0.019	1.488	1.157
30		0.283	0.292	1.384	1.390
31		-0.486	-0.483	1.419	1.455
32		-0.241	-0.251	1.505	1.516
33		0.019	0.008	1.568	1.613
34		0.261	0.245	1.422	1.624
35		0.506	0.491	1.081	1.113
36		-0.248	-0.277	1.384	1.398
37		0.016	-0.002	1.057	1.139
38		0.273	0.269	1.213	1.394
39		0.019	-0.004	1.095	1.171

表 2 構造ベクトル  
Table 2 Structure vectors.

	$\eta_1$	$\eta_2$
1	0.999	0.001
2	0.999	-0.006
3	0.001	0.999
4	0.007	0.998
$r$	1.998	1.996

$\phi_1, \phi_2$  によって変換されたパターン  $y_i = (\phi_1, \phi_2)^T x_i$  を示している。 $\phi_1$  による特徴量  $y_1$  をみると  $C_1$  と  $C_2$  の特徴がよく抽出され、結果として  $C_3$  ともよく類別されている。 $\phi_2$  による  $y_2$  をみると  $C_3$  の特徴がよく抽出され、 $C_1, C_2$  のそれとよく類別されている。図 3

表 3 単純化ベクトル  
Table 3 Simplistic vectors.

	$\phi_1$	$\phi_2$
1	0.703	-0.005
2	0.710	-0.000
3	0.003	0.713
4	-0.000	0.700

表 4 単純構造変換による特徴量  
Table 4 Features by simple-structure transformation.

Class		$v_1$	$v_2$
$C_1$	1	1.721	-0.937
	2	1.708	-1.503
	3	1.640	-0.972
	4	1.672	-0.371
	5	1.698	-2.057
	6	1.747	-1.492
	7	1.720	-0.927
	8	1.740	-0.405
	9	1.748	0.138
	10	1.702	-1.462
	11	1.689	-0.950
	12	1.675	-0.418
	13	1.667	-0.945
$C_2$	14	-1.659	-0.943
	15	-1.681	-1.483
	16	-1.696	-0.940
	17	-1.675	-0.431
	18	-1.676	-1.993
	19	-1.691	-1.419
	20	-1.738	-0.935
	21	-1.700	-0.396
	22	-1.729	0.096
	23	-1.750	-1.455
	24	-1.727	-0.928
	25	-1.829	-0.411
	26	-1.793	-0.938
	$C_3$	27	0.024
28		-0.333	1.801
29		-0.002	1.872
30		0.411	1.959
31		-0.680	2.034
32		-0.343	2.137
33		0.024	2.248
34		0.362	2.150
35		0.708	1.548
36		-0.366	1.968
37		0.013	1.552
38		0.387	1.840
39		0.014	1.601

表 5 正規化された固有ベクトル  
Table 5 Normalied eigen vectors.

	$\phi_1$	$\phi_2$
1	0.568	-0.419
2	0.565	-0.424
3	0.420	0.568
4	0.424	0.565
固有値	2.001	1.993

表 6 K-L 変換による特徴量  
Table 6 Feature by Karhunen-Loève transform.

Class		$v_1$	$v_2$
$C_1$	1	0.827	-1.775
	2	0.479	-2.222
	3	0.740	-1.755
	4	1.124	-1.291
	5	0.141	-2.662
	6	0.517	-2.237
	7	0.831	-1.767
	8	1.158	-1.359
	9	1.489	-0.926
	10	0.498	-2.186
	11	0.793	-1.766
	12	1.009	-1.330
	13	0.779	-1.749
$C_2$	14	-1.898	0.227
	15	-2.236	-0.194
	16	-1.925	0.251
	17	-1.604	0.648
	18	-2.535	-0.606
	19	-2.206	-0.135
	20	-1.955	0.281
	21	-1.604	0.691
	22	-1.334	1.104
	23	-2.275	-0.130
	24	-1.942	0.279
	25	-1.715	0.755
	26	-2.001	0.311
	$C_3$	27	1.075
28		0.801	1.645
29		1.110	1.504
30		1.497	1.330
31		0.664	2.039
32		0.996	1.921
33		1.358	1.792
34		1.573	1.513
35		1.491	0.823
36		0.877	1.799
37		0.935	1.259
38		1.408	1.249
39		0.965	1.278

は  $v_1$  に対するヒストグラムで三つのクラスがよく類別されている有様を明示したものである。

次に同じデータをもちい、パターンクラス生起確率を  $P_1=P_2=P_3=1/3$  にとって K-L 変換を行ってみた。表 5 の  $\phi_1, \phi_2$  は得られた固有ベクトルを正規化し

たものである。 $\phi_1, \phi_2$  の方向に抽出される特徴量が全パターンクラスに相関の強いものとなることがよくわかる。表 6 は  $\phi_1, \phi_2$  による特徴量である。 $\phi_1$  による  $v_1$  をみると  $C_2$  は他の  $C_1, C_3$  とよく類別されているが、 $C_1$  と  $C_3$  はひどく重なりがみられる。 $\phi_2$  による

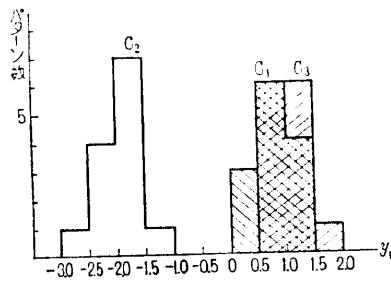


図4 K-L変換による特徴量  $y_1$  のヒストグラム  
Fig. 4 Histogram for feature  $y_1$  by Karhunen-Loève transform.

$y_2$  をみると  $C_1$  が他の  $C_2, C_3$  とよく類別されているが、 $C_2, C_3$  は重なりがみられる。図4は  $y_1$  に対するヒストグラムである。

この例では三つのクラスを類別するのに単純構造変換では1次元部分空間にパターンを射影するだけでよいが、K-L変換では少なくとも2次元の部分空間に射影する必要がある。この実験結果からもわかるように単純構造変換が形成する部分空間の各軸は少なくとも一つ(例では  $\phi_1$  方向の軸は二つ)のクラスのもっとも重要な情報を担うことになる。それに対しK-L変換での軸は全クラスの情報を担っている。K-L変換はクラス類別のための特徴を抽出するのに必ずしも最適ではないことがわかる。一般に  $K$  個のパターンクラスが分布し、各クラスの特徴がそれぞれ別個の変量グループによって観測されている場合、高々  $K$  個の単純化ベクトルがつくる部分空間にパターンを射影するのがもっとも適当といえる。

## 6. むすび

パターンクラスの特徴が主に一部の変量グループによって観測されているとわかっているものの、どのようなグループになっているのか不明のとき、単純構造基準を最大にすることによってグループを見つけだし、同時にその変量グループにもっとも相関の強い特徴量を抽出する単純化ベクトルを見いだす方法について述べた。そしてこれらのベクトルがつくる部分空間にパターンを射影するならばクラスを類別する重要な特徴が得られることを実験により確かめ、従来のK-L変換よりも有用であることを示した。本方法のもっとも大きな特徴は単純化ベクトルがつくる部分空間の各軸は少なくとも一つのクラスの重要な情報を担うという点であろう。一般に  $K$  個のクラスならば高々  $K$  個の単純化ベクトルでよい。本方法は高次元多クラス問

題においてクラスの特徴が主として一部の変量グループによって観測されている場合の一つの特徴抽出法として有効に働くものと思われる。今後実在する高次元パターンに対し本方法の有用性を確かめていきたい。

最後に日ごろより有益な討論をしていただく当研究室の金岡泰保氏に深謝する。

## 参考文献

- 1) Watanabe, S.: Karhunen-Loève Expansion and Factor Analysis, Trans. 4th Prague Conf. on Information Theory, pp. 635-660 (1965).
- 2) Chien, Y. T. and Fu, K. F.: On the Generalized Karhunen-Loève Expansion, IEEE Trans. Inform. Theory, IT-13, pp. 518-520 (1967).
- 3) Fukunaga, K. and Koontz, W. L. G.: Application of the Karhunen-Loève Expansion to Selection and Ordering, IEEE Trans. Comput., C-19, pp. 311-318 (1970).
- 4) Foley, D. H. and Sammon, J. W. Jr.: An Optimal Set of Discriminant Vectors, IEEE Trans. Comput., C-24, pp. 281-289 (1975).
- 5) 芝 祐順: 因子分析法, p. 424, 東京大学出版会 (1975).
- 6) 遠山 啓: 行列論, p. 266, 共立出版 (1975).
- 7) 富田, 野口, 大泉: Karhunen-Loève 直交系による教師なしの類別, 電子通信学会論文誌 C, Vol. 54-C, No. 8, pp. 767-774 (1971).

## 付 録

[1] 一般にパターン認識系においては観測、特徴抽出、識別の各処理系がシステムとして互に関連を保ち、一つの処理系を構成する。しかし、対象とするパターン集合について事前に何らかの知識を与えられていない場合は、いわゆる教師なしの構成となるため、その処理機構は非常に複雑で、かつ実用的でない<sup>7)</sup>。しかし、対象のパターン集合があらかじめ何らかの知識、例えば  $n$  次元空間に分布を構成しているとか、パターンの塊がいくつあるとか等が与えられていれば、その知識を基にパターン処理系を構成する設計方法を組み立て、さらに一般的な場合へと拡張するのが妥当である。本論文はこのような立場から与えられたパターン集合はあらかじめ何らかの知識、すなわち  $n$  次元空間にパターンの塊(これをパターンクラス、あるいは単にクラスとよぶ)が  $K$  個存在すると仮定し、これらの塊がパターン集合全体を特徴づけている、すなわち塊の一つの特徴とみなして、それを抽出することを考察する。

ところで各クラスの相違は各クラスに属するパター

ンの観測値, すなわちパターンベクトルを構成する  $n$  個の変量の観測値に起因するのは明らかである. しかし,  $n$  個の変量のすべてが同程度によくクラス間の相違を観測しているとはかぎらない. 各変量の観測値はクラス間の相違に異なった寄与を与えると考えるのが自然であろう. 換言すれば各クラスにおいては他クラスとの相違をよく観測している変量とそうでない変量が存在する. 本論文ではこのような観点から各クラスにおける他クラスとの相違特徴, すなわちクラス特有の特徴をよく観測している変量を因子分析法における単純構造の概念を導入して見つけ, それらの変量にもっとも相関の強い特徴を抽出する方法について述べている.

[2] いま  $\mathbf{p}_2^*$  を  $\mathbf{X}$  から任意ベクトル  $\xi_2^*$  をもちいて得られるベクトルとする. すなわち (4) 式の  $\mathbf{p}_1$  と同様に

$$\mathbf{p}_2^* = \mathbf{X}\xi_2^*. \quad (\text{A1})$$

ただし,  $\mathbf{p}_2^*$  と  $\mathbf{p}_1$  の間には  $\mathbf{p}_2^* = c\mathbf{p}_1$  を満たす定数  $c$  は  $c=0$  のみとする.  $\mathbf{p}_2$  を

$$\mathbf{p}_2 = \mathbf{p}_2^* - c\mathbf{p}_1 \quad (\text{A2})$$

とおき,  $\mathbf{p}_1^T \mathbf{p}_2 = 0$  を満足するように  $c$  を定めると

$$c = \mathbf{p}_1^T \mathbf{p}_2^* / \mathbf{p}_1^T \mathbf{p}_1 = \xi_1^T \mathbf{R}\xi_2^* / \xi_1^T \mathbf{R}\xi_1 \quad (\text{A3})$$

となる. したがって

$$\begin{aligned} \mathbf{p}_2 &= \mathbf{X}\xi_2^* - (\xi_1^T \mathbf{R}\xi_2^* / \xi_1^T \mathbf{R}\xi_1) \mathbf{X}\xi_1 \\ &= \mathbf{X}(\mathbf{I} - \xi_1 \xi_1^T \mathbf{R} / \xi_1^T \mathbf{R}\xi_1) \xi_2^* \\ &= \mathbf{X}(\mathbf{I} - \xi_1 \eta_1^T \mathbf{S} / \sqrt{\xi_1^T \mathbf{R}\xi_1}) \xi_2^* \\ &= \mathbf{X}_1 \xi_2^* \end{aligned}$$

となる (証明終).

[3]  $\mathbf{p}_1$  および  $\mathbf{p}_2^*$  に準じて

$$\begin{aligned} \mathbf{p}_{i-1} &= \mathbf{X}\xi_{i-1} = \mathbf{X}_{i-2} \xi_{i-1}^* \\ \mathbf{p}_i^* &= \mathbf{X}_{i-2} \xi_i^* \end{aligned}$$

とおく. ただし,  $\mathbf{p}_i^* = c\mathbf{p}_{i-1}$  が成立する定数  $c$  は  $c=0$  だけとする.

$$\mathbf{p}_i = \mathbf{p}_i^* - c\mathbf{p}_{i-1}$$

とおき,  $\mathbf{p}_{i-1}^T \mathbf{p}_i = 0$  を満足するように  $c$  を定めると

$$\begin{aligned} c &= \mathbf{p}_{i-1}^T \mathbf{p}_i^* / \mathbf{p}_{i-1}^T \mathbf{p}_{i-1} \\ &= \xi_{i-1}^{*T} \mathbf{R}_{i-2} \xi_i^* / \xi_{i-1}^{*T} \mathbf{R}_{i-2} \xi_{i-1}^* \end{aligned}$$

となる. したがって

$$\begin{aligned} \mathbf{p}_i &= \mathbf{p}_i^* - (\xi_{i-1}^{*T} \mathbf{R}_{i-2} \xi_i^* / \xi_{i-1}^{*T} \mathbf{R}_{i-2} \xi_{i-1}^*) \mathbf{p}_{i-1} \\ &= (\mathbf{X}_{i-2} \\ &\quad - \mathbf{X}_{i-2} \xi_{i-1}^* \xi_{i-1}^{*T} \mathbf{R}_{i-2} / \xi_{i-1}^{*T} \mathbf{R}_{i-2} \xi_{i-1}^*) \xi_i^* \\ &= (\mathbf{X}_{i-2} - \mathbf{X}\xi_{i-1} \eta_{i-1}^T \mathbf{S} / \sqrt{\xi_{i-1}^T \mathbf{R}\xi_{i-1}}) \xi_i^* \end{aligned}$$

$\mathbf{X}_{i-2}$  に (20) 式の  $\mathbf{X}_{i-2}$  の場合を代入すれば

$$\mathbf{p}_i = \mathbf{X} \left( \mathbf{I} - \sum_{j=1}^{i-1} \xi_j \eta_j^T \mathbf{S} / \sqrt{\xi_j^T \mathbf{R}\xi_j} \right) \xi_i^*$$

となる (証明終).

[4] 以下の証明の中で式を簡潔にするために (32) 式を  $n$  倍したものを

$$W = nV = nE[\eta_j^4] - n(E[\eta_j^2])^2 \quad (\text{B1})$$

と表わすことにする.  $\mathbf{p}_1$  を単純構造にするため  $W$  を

$$W_1 = \sum_{j=1}^n \eta_{j1}^4 - \left( \sum_{j=1}^n \eta_{j1}^2 \right)^2 / n \quad (\text{B2})$$

とおき, これを最大にする  $\eta_1$  を  $\mathbf{R}$  の成分  $r_{jk}$  が

$$r_{jk} = s_{js} s_{sk} \sum_{p=1}^m \eta_{jp} \eta_{kp} \quad (\text{B3})$$

となる条件のもとで求めてみる. そのために

$$4U = W_1 - 2 \sum_{j=1}^n \sum_{k=1}^n \mu_{jk} \left( \sum_{p=1}^m \eta_{jp} \eta_{kp} - r_{jk} / s_{js} s_{sk} \right) \quad (\text{B4})$$

のような値  $U$  を考え,  $U$  を各  $\eta_{jp}$  ( $p=1, \dots, m$ ) について偏微分し, 結果をすべて 0 とおく. ここで,  $\mu_{jk}$  はラグランジュの乗数である. まず  $\partial U / \partial \eta_{j1} = 0$  を実行すると

$$\eta_{j1}^3 - \eta_{j1} \sum_{k=1}^n \eta_{k1}^2 / n = \sum_{k=1}^n \mu_{jk} \eta_{k1} \left( = \sum_{k=1}^n \mu_{kj} \eta_{k1} \right). \quad (\text{B5})$$

$\partial U / \partial \eta_{jp} = 0$  ( $p \neq 1$ ) を実行すると

$$\sum_{k=1}^n \mu_{jk} \eta_{kp} = 0 \quad (\text{B6})$$

を得る. (B5) と (B6) を一つの式にまとめると

$$\delta_{1p} \left( \eta_{jp}^3 - \eta_{jp} \sum_{k=1}^n \eta_{k1}^2 / n \right) = \sum_{k=1}^n \mu_{jk} \eta_{kp} \quad (\text{B7})$$

ここに,  $\delta_{1p}$  は  $p=1$  のとき 1,  $p \neq 1$  のとき 0 の値をとる. 両辺に  $\eta_{j1}$  を乗じて  $j$  について和をとると

$$\begin{aligned} \delta_{1p} \left( \sum_{j=1}^n \eta_{j1}^4 - \sum_{j=1}^n \eta_{j1}^2 \sum_{k=1}^n \eta_{k1}^2 / n \right) \\ = \sum_{j=1}^n \sum_{k=1}^n \mu_{jk} \eta_{j1} \eta_{kp} \end{aligned}$$

となり, これは

$$\delta_{1p} W_1 = \sum_{k=1}^n \eta_{kp} \sum_{j=1}^n \mu_{jk} \eta_{j1}$$

と書ける. 右辺の  $\sum_{j=1}^n \mu_{jk} \eta_{j1}$  は (B5) の  $j$  と  $k$  を交換したものにほかならない. したがって

$$\delta_{1p} W_1 = \sum_{k=1}^n \eta_{kp} \zeta_{k1}.$$

ここに,  $\zeta_{k1} = \eta_{k1}^3 - \eta_{k1} \sum_{k=1}^n \eta_{k1}^2 / n$  である.



さらに、両辺に  $\eta_{1p}$  を乗じて  $p$  について和をとると

$$\begin{aligned} \sum_{p=1}^m \delta_{1p} \eta_{1p} W_1 &= \sum_{k=1}^n \zeta_{k1} \sum_{p=1}^m \eta_{1p} \eta_{kp} \\ &= \sum_{k=1}^n \zeta_{k1} r_{1k} / s_{1s_k} \end{aligned}$$

となる。したがって

$$\eta_{11} W_1 = \sum_{k=1}^n \zeta_{k1} r_{1k} / s_{1s_k}$$

となり、まとめて行列とベクトルで表わすと

$$W_1 \eta_1 = S^{-1} R S^{-1} \zeta_1 \tag{B8}$$

ここに、 $\zeta_1 = \eta_1^{(3)} - \eta_1^T \eta_1 \eta_1 / n$ ,

$$\eta_1^{(3)} = (\eta_{11}^3, \eta_{21}^3, \dots, \eta_{m1}^3)^T.$$

一方、(B8)の両辺に  $\zeta_1^T$  を乗じると

$$W_1 \zeta_1^T \eta_1 = \zeta_1^T S^{-1} R S^{-1} \zeta_1.$$

また、 $\zeta_1^T \eta_1 = W_1$  より

$$W_1^2 = \zeta_1^T S^{-1} R S^{-1} \zeta_1.$$

したがって

$$W_1 = \sqrt{\zeta_1^T S^{-1} R S^{-1} \zeta_1}.$$

結局 (B8) は

$$\eta_1 = S^{-1} R S^{-1} \zeta_1 / \sqrt{\zeta_1^T S^{-1} R S^{-1} \zeta_1}.$$

これと(7)式  $\eta_1 = S^{-1} R \xi_1 / \sqrt{\xi_1^T R \xi_1}$  を比較して

$$\xi_1 = S^{-1} \zeta_1$$

を得る (証明終).

[5]  $p_1$  と直交し、かつ単純構造をなす  $p_2$  を得るために

$$W_2 = \sum_{j=1}^n \eta_{j2}^4 - \left( \sum_{j=1}^n \eta_{j2}^2 \right)^2 / n \tag{C1}$$

とおき、これを最大にする  $\eta_2$  を  $R_1$  の成分  $r_{jk}$  が

$$r_{jk} = r_{jk} - s_{js_k} \eta_{j1} \eta_{k1} = s_{js_k} \sum_{p=2}^m \eta_{jp} \eta_{kp} \tag{C2}$$

となる条件のもとで求めればよい。(B4)に対応する

$$4U = W_2 - 2 \sum_{j=1}^n \sum_{k=1}^n \mu_{jk} \left( \sum_{p=2}^m \eta_{jp} \eta_{kp} - r_{jk} / s_{js_k} \right) \tag{C3}$$

なる値  $U$  を考え、これを各  $\eta_{jp}$  ( $p=2, \dots, m$ ) について偏微分し、結果をすべて0とおく。以下前と同様の手法で進めれば

$$\eta_2 = S^{-1} R_1 S^{-1} \zeta_2 / \sqrt{\zeta_2^T S^{-1} R_1 S^{-1} \zeta_2}$$

を得る。これと(15)式  $\eta_2 = S^{-1} R_1 \xi_2^* / \sqrt{\xi_2^{*T} R_1 \xi_2^*}$  を比較して、 $\xi_2^* = S^{-1} \zeta_2$  が得られ、(12)式より

$$\xi_2 = (I - \xi_1 \eta_1^T S / \sqrt{\xi_1^T R \xi_1}) S^{-1} \zeta_2$$

を得る (証明終)。以下同様にして証明でき、一般に  $p_1, p_2, \dots, p_{i-1}$  と直交し、かつ単純構造をなす  $p_i$  を抽出する  $\xi_i$  を順次得る。

(昭和53年8月25日受付)

(昭和54年5月17日採録)

## 因子分析法による特徴抽出†

岡田敏彦†† 富田真吾††

本論文は因子分析法における単純構造の概念を基にし、パターンの特徴抽出問題を論じている。すなわち  $n$  次元空間上のパターンクラスは、クラス間の相違が各クラスのパターンの  $n$  個の変量の観測値に依存するが、常に  $n$  個の変量のすべてが同程度にクラス間の相違に影響を及ぼしているとはかぎらない。このため、各クラスにおいては他のクラスとの相違に大きな役割を果たしている変量グループとそうではない変量グループが存在する。この事実に着目し、クラス間の相違に大きな役割を果たしている変量グループを明らかにする目的で因子分析法の単純構造化の手法を用いることにより、これらの変量グループにもっとも相関の強い特徴を抽出する単純構造系を提案している。単純構造系の座標軸はそれぞれ各クラスに対応し、クラスの類別に重要な情報を担い、クラスが  $K$  個分布しているなら、高々  $K$  個の座標軸よりなる系として構成される。さらに具体的なパターンをもちいて計算機シミュレーションを行い、従来の Karhunen-Loève 系との比較検討し、本方法の有用性について述べている。

## 1. ま え が き

パターンの特徴抽出法として従来より種々の方法が提案されてきている。パターンの特徴抽出の目的はパターンの膨大な観測値、たとえば一つの波形や画像をデジタル化するとしばしば膨大な量の観測値集合となるが、これらのものからパターンを識別する情報を十分含むより小さな特徴量集合に、もとのパターン観測値を写像する変換を見つけることである。よく知られた方法の一つに Karhunen-Loève 変換 (K-L 変換) による方法、すなわち、正規直交固有ベクトル変換<sup>1),2)</sup>がある。この方法はパターン自己相関行列の固有ベクトルに対応する固有値の大きい順に選ばれ、選ばれた固有ベクトルがつくる部分空間にもとのパターンを射影するものである。この部分空間には各パターンにもっとも相関の強い特徴量が抽出され、この点に関しては最適な空間であるが、パターンを類別するための特徴抽出空間としては必ずしも最適ではない。この欠点を補うため 1970 年 Fukunaga-Koontz<sup>3)</sup> は修正された K-L 変換法を提案した。すなわち 2 クラス類別問題において二つのクラスの重みつきパターン自己相関行列をそれぞれ  $R_1, R_2$  とし、これらの和  $R_0 = R_1 + R_2$  に対してある予備変換  $PR_0P^T = I$  ( $P$  は変換行列,  $I$  は単位行列,  $T$  は転置を示す) を行うとき、変換された両クラスの自己相関行列  $PR_1P^T, PR_2P^T$  は同じ固有ベクトルと固有値をもつ。クラス

1 に対する固有値が  $1 \geq \lambda_1^{(1)} \geq \lambda_2^{(2)} \geq \dots \geq \lambda_n^{(1)} \geq 0$  のように並べられるとすれば、クラス 2 に対する固有値は  $\lambda_i^{(2)} = 1 - \lambda_i^{(1)}$  となる。したがって  $\lambda_i^{(1)}$  に対応する固有ベクトルはクラス 1 のもっとも重要な特徴を抽出するがクラス 2 に対してはもっとも重要でない特徴を抽出する。逆に  $\lambda_n^{(2)} (= 1 - \lambda_n^{(1)})$  に対応する固有ベクトルはクラス 2 のもっとも重要な特徴を抽出するがクラス 1 に対してはもっとも重要でない特徴を抽出する。すなわち両クラスは共通に重要な特徴は共有できなくなる。Fukunaga らは  $|\lambda_i^{(1)} - 0.5|$  の値の大きい順に対応する固有ベクトルを選び、選ばれた固有ベクトルがつくる部分空間にパターンを射影することを提案している。これは非常にすぐれた方法であるが一般の多クラス問題に拡張できないのが欠点である。1975 年 Foley-Sammon<sup>4)</sup> は 2 クラス問題に対し二つのクラス間の距離をできるだけ大きく、同時にクラス内のパターンの分散をできるだけ小さくする変換ベクトルを見つけ、これらのベクトルがつくる部分空間にパターンを射影する方法を提案している。そして Fukunaga-Koontz 変換法ではうまく類別できない例についても良好に類別することを示している。しかしやはり 2 クラス問題に限定され一般性に欠けている。1965 年 S. Watanabe<sup>1)</sup> は K-L 系のパターン認識への応用にもとづき一般的な因子分析法による特徴抽出への応用について示唆を与えている。本論文ではこの示唆をもとに 3 個以上の多クラス問題について考察した。すなわち  $n$  次元空間に  $K$  個のパターンクラス  $C_i$  ( $i=1, \dots, K$ ) が分布しているとしよう。各クラスはそれぞれ特有の特徴を有していて、これらの特徴は  $n$  個の変量のうち主として一部の変量グループによって観測され

† Feature Extraction of Patterns by Factor Analysis by TOSHIHIKO OKADA and SHINGO TOMITA (Faculty of Engineering, Yamaguchi University).

†† 山口大学工学部電子工学科

ているものとする. すなわちあるクラス  $C_i$  の特徴は主として  $n_i$  個の変量が, またあるクラス  $C_j$  の特徴は別の  $n_j$  個の変量が主に観測している場合である (付録 [1] を参照). しかしこれらの変量グループは事前にはわかっていないとしよう. このようなとき因子分析法における単純構造化の概念を導入し, 変量のグループ分けを行い, 各変量グループにもっとも相関の強い方向ベクトルを見つけだし, それらのベクトルがつくる部分空間にパターンを射影するならばその空間は  $K$  個のパターンクラスの類別に有用な特徴を抽出するであろう. しかしここで問題は因子分析法ではパターンを各変量について平均が 0, 分散が 1 となるように基準化することである. 平均を 0 にする操作は単に空間の原点の移動するだけであり, パターンの分布形状になんの影響も与えないが, 分散を 1 にする操作はパターンの分布形状を変えてしまう. これはパターンの特徴抽出を考えるうえで非常に問題になる. したがって因子分析法の理論をそのまま導入するには問題がある. そこで本論文では理論を展開するうえで式を簡潔に表現するためと, パターンの分布形状に影響を与えないとの理由から各変量について平均が 0 になるように前処理するが, 分散についてはそのままにする. そのうえで因子分析法の単純構造化の理論を特徴抽出への応用のために再構成する.

2章で数学的準備と特徴抽出機構の基本原則を, 3章で特徴抽出のための単純構造化変換理論を述べる. そして4章で K-L 変換との相違について述べ, 5章で本変換法の有用性を確かめるため計算機実験の結果を示し, 同じデータを使って行った K-L 変換の結果と比較検討をする.

## 2. 準備

$n$  次元空間に  $N$  個のパターン集合  $\{x_1, x_2, \dots, x_N\}$  が与えられ, 原点は  $\sum_{i=1}^N x_i = 0$  となるようにとられているものとしよう. すなわち

$$\sum_{i=1}^N x_{ij} = 0 \quad (j=1, \dots, n) \quad (1)$$

である. ここに,  $x_i = (x_{i1}, x_{i2}, \dots, x_{in})^T$ ,  $x_{ij}$  はパターン  $x_i$  の変量  $j$  の観測値を表わし実数である.

パターン行列を

$$X = (x_1, x_2, \dots, x_N)^T, \quad (2)$$

パターン自己相関行列を

$$R = E[x_i x_i^T] = X^T X / N \quad (3)$$

として表わす. パターン行列  $X$  よりベクトル  $\xi_1$  によって抽出される特徴量を成分とするベクトルを

$$p_1 = X \xi_1 \quad (4)$$

と表わす.  $p_1$  の成分  $p_{1i} (i=1, \dots, N)$  はパターン  $x_i$  より抽出された特徴量を示す.  $p_1$  の成分の平均は 0, 分散は  $\xi_1^T R \xi_1$  となる.  $p_1$  と各変量との相関係数を成分とするベクトル  $\eta_1$  は,  $p_1$  を基準化したものを

$$q_1 = p_1 / \sqrt{\xi_1^T R \xi_1} \quad (5)$$

と表わし,  $X$  の各列について, すなわち各変量について基準化したものを

$$X_s = X S^{-1} \quad (6)$$

と表わせば

$$\begin{aligned} \eta_1 &= X_s^T q_1 / N \\ &= S^{-1} R \xi_1 / \sqrt{\xi_1^T R \xi_1} \end{aligned} \quad (7)$$

となる. ここに行列  $S$  は各変量の標準偏差を対角成分とするもので  $S = \text{diag}(s_1, s_2, \dots, s_n)$  であり  $s_i$  は変量  $i$  の標準偏差である. なお  $s_i \neq 0 (i=1, \dots, n)$  とする.  $\eta_1$  は特徴量の構造を表わし, 因子分析法における因子負荷ベクトルに相当するが, ここでは構造ベクトルとよぶことにする.  $\eta_1$  の  $i$  番目の成分  $\eta_{1i}$  は  $p_1$  と変量  $i$  との相関係数を示す.

$p_1$  はパターン間のある種の特徴差を, ひいてはクラス間の特徴差を表わしうるが, さらに  $p_1$  では表わしえなかった別種の特徴差を表わしたいことがある. それを  $p_2$  で表わしたいとき,  $p_2$  は  $p_1$  とはまったく相関のない独立したものにすることがある. すなわち  $p_2$  を  $p_1$  と直交化して

$$p_1^T p_2 = 0$$

となるようにしなければならない. このような  $p_2$  は

$$p_2 = X(I - \xi_1 \eta_1^T S / \sqrt{\xi_1^T R \xi_1}) \xi_2^* \quad (8)$$

$$= X \xi_2^* \quad (9)$$

$$= X \xi_2 \quad (10)$$

として得られる (証明は付録 [2]).

ここに,

$$X_1 = X(I - \xi_1 \eta_1^T S / \sqrt{\xi_1^T R \xi_1}) \quad (11)$$

$$\xi_2 = (I - \xi_1 \eta_1^T S / \sqrt{\xi_1^T R \xi_1}) \xi_2^* \quad (12)$$

で,  $\xi_2^*$  は任意ベクトルである.

この  $X_1$  をもちいて

$$R_1 = X_1^T X_1 / N = R - S \eta_1 \eta_1^T S \quad (13)$$

とおく.  $p_2$  の成分の平均は 0, 分散は  $\xi_2^{*T} R_1 \xi_2^*$  となるので  $p_2$  と各変量との相関係数を成分とするベクトル  $\eta_2$  は,  $p_2$  を基準化して

$$q_2 = p_2 / \sqrt{\xi_2^{*T} R_1 \xi_2^*} \quad (14)$$

とおけば

$$\begin{aligned} \eta_2 &= \mathbf{X}_1^T \mathbf{q}_2 / N \\ &= \mathbf{S}^{-1} \mathbf{R}_1 \xi_2^* / \sqrt{\xi_2^{*T} \mathbf{R}_1 \xi_2^*} \\ &= \mathbf{S}^{-1} \mathbf{R}_2 \xi_2 / \sqrt{\xi_2^T \mathbf{R}_2 \xi_2} \end{aligned} \quad (15)$$

と表わされる。η<sub>2</sub> の *i* 番目の成分 η<sub>2*i*</sub> は **p**<sub>2</sub> と変数 *i* との相関係数を示す。

一般に **p**<sub>1</sub>, **p**<sub>2</sub>, ..., **p**<sub>*i*-1</sub> と直交する **p**<sub>*i*</sub> は

$$\mathbf{p}_i = \mathbf{X} \left( \mathbf{I} - \sum_{j=1}^{i-1} \xi_j \eta_j^T \mathbf{S} / \sqrt{\xi_j^T \mathbf{R}_j \xi_j} \right) \xi_i^* \quad (17)$$

$$= \mathbf{X}_{i-1} \xi_i^* \quad (18)$$

$$= \mathbf{X} \xi_i \quad (19)$$

となる (証明は付録 [3])。

ここに、

$$\mathbf{X}_{i-1} = \mathbf{X} \left( \mathbf{I} - \sum_{j=1}^{i-1} \xi_j \eta_j^T \mathbf{S} / \sqrt{\xi_j^T \mathbf{R}_j \xi_j} \right) \quad (20)$$

$$\xi_i = \left( \mathbf{I} - \sum_{j=1}^{i-1} \xi_j \eta_j^T \mathbf{S} / \sqrt{\xi_j^T \mathbf{R}_j \xi_j} \right) \xi_i^* \quad (21)$$

で、ξ<sub>*i*</sub><sup>\*</sup> は任意ベクトルである。また η<sub>*i*</sub> は

$$\eta_i = \mathbf{S}^{-1} \mathbf{R}_{i-1} \xi_i^* / \sqrt{\xi_i^{*T} \mathbf{R}_{i-1} \xi_i^*} \quad (22)$$

$$= \mathbf{S}^{-1} \mathbf{R}_i \xi_i / \sqrt{\xi_i^T \mathbf{R}_i \xi_i} \quad (23)$$

ここに、

$$\begin{aligned} \mathbf{R}_{i-1} &= \mathbf{X}_{i-1}^T \mathbf{X}_{i-1} / N \\ &= \mathbf{R} - \mathbf{S} \sum_{j=1}^{i-1} \eta_j \eta_j^T \mathbf{S} \end{aligned} \quad (24)$$

である。ところで(20)式は

$$\mathbf{X}_{i-1} = \mathbf{X} - \sum_{j=1}^{i-1} \mathbf{p}_j \eta_j^T \mathbf{S} / \sqrt{\xi_j^T \mathbf{R}_j \xi_j}$$

と書くことができ、さらに

$$\mathbf{r}_j^T = \eta_j^T \mathbf{S} / \sqrt{\xi_j^T \mathbf{R}_j \xi_j}$$

とおけば

$$\mathbf{X}_{i-1} = \mathbf{X} - \sum_{j=1}^{i-1} \mathbf{p}_j \mathbf{r}_j^T$$

と書ける。したがって、**X** は

$$\mathbf{X} = \mathbf{p}_1 \mathbf{r}_1^T + \mathbf{p}_2 \mathbf{r}_2^T + \dots + \mathbf{p}_m \mathbf{r}_m^T + \dots$$

と表わされる。**X** の階数を *m* とすれば **X** は *m* 個の互いに直交するベクトルで表わすことができ<sup>6)</sup>、いま **p**<sub>*i*</sub> と **p**<sub>*j*</sub> (*i* ≠ *j*) は互いに直交するベクトルであるので、結局 **X** は

$$\mathbf{X} = \mathbf{p}_1 \mathbf{r}_1^T + \mathbf{p}_2 \mathbf{r}_2^T + \dots + \mathbf{p}_m \mathbf{r}_m^T \quad (25)$$

と表わされる。**X**<sup>T</sup> の階数も *m* であるから **R** = **X**<sup>T</sup> **X** / *N* の階数は *m* となる。したがって **R** は

$$\begin{aligned} \mathbf{R} &= \mathbf{X}^T \mathbf{X} / N \\ &= \mathbf{S} \sum_{i=1}^m \eta_i \eta_i^T \mathbf{S} \end{aligned} \quad (26)$$

と表わされる。

一方特徴抽出ベクトル ξ<sub>*i*</sub> は方向のみが重要であるから正規化したものを

$$\phi_i = \xi_i / \|\xi_i\| \quad (\|\phi_i\|^2 = \phi_i^T \phi_i = 1) \quad (27)$$

とおく。さらに φ<sub>*i*</sub> に対応した η<sub>*i*</sub> をもちいて

$$\gamma_i = \eta_i^T \phi_i \quad (28)$$

を定義する。η<sub>*i*</sub> の成分は抽出された特徴量と各変数との相関係数であり、その二乗和を γ<sub>*i*</sub> は示すので、いま γ<sub>*i*</sub> が大きいことは γ<sub>*i*</sub> に対応した φ<sub>*i*</sub> によって抽出される特徴量が変数と強い相関をもつことになる。すべての変数と、あるいは一部の変数と強い相関をもつかは φ<sub>*i*</sub> の設計にかかってくる。なぜなら φ<sub>*i*</sub> は(21)式からわかるように任意ベクトル ξ<sub>*i*</sub><sup>\*</sup> に依存するからである。したがって、γ<sub>*i*</sub> の大きい順に対応する φ<sub>*i*</sub> を選び、それらがつくる部分空間にパターンを射影するならば、すべての変数、あるいは一部の變数と相関の強い特徴が抽出されることになる。選ばれた φ<sub>*i*</sub> を φ<sub>1</sub>, φ<sub>2</sub>, ..., φ<sub>*i*</sub> とすれば

$$\Phi = (\phi_1, \phi_2, \dots, \phi_i)^T \quad (29)$$

は

$$\mathbf{y} = (y_1, y_2, \dots, y_i)^T = \Phi \mathbf{x} \quad (30)$$

のような変換を行うことになり

$$y_i = \phi_i^T \mathbf{x} \quad (31)$$

はパターン **x** の新しい空間における *i* 番目の特徴量となる。この特徴抽出機構を図1に示す。

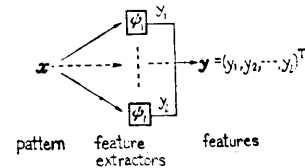


図1 特徴抽出機構

Fig. 1 Feature extraction mechanism.

また φ<sub>*i*</sub> は任意ベクトル ξ<sub>*i*</sub><sup>\*</sup> に依存していることから、φ<sub>*i*</sub>(ξ<sub>*i*</sub><sup>\*</sup>) を可変することにより特徴量構造を制御することが可能となる。ここではパターンの類別のための特徴量構造にするにはどのように φ<sub>*i*</sub>(ξ<sub>*i*</sub><sup>\*</sup>) を設計すればよいかについて考え、次の3章で詳しく述べる。

図2は φ<sub>*i*</sub>(ξ<sub>*i*</sub><sup>\*</sup>) による y<sub>*i*</sub> の制御を図示したものである。

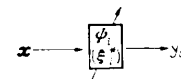


図2 可変 φ<sub>*i*</sub>(ξ<sub>*i*</sub><sup>\*</sup>) による y<sub>*i*</sub> の制御

Fig. 2 Control of y<sub>*i*</sub> by a variable φ<sub>*i*</sub>(ξ<sub>*i*</sub><sup>\*</sup>).

### 3. 単純構造変換

$n$ 次元空間に  $K$  個のパターンクラス  $C_i$  ( $i=1, \dots, K$ ) が分布し、各クラスの特徴は主に一部の変量グループによって観測されているものとしよう (付録 [1] 参照). すなわち、あるクラス  $C_i$  の特徴は  $n$  個の変量のうち主に  $n_i$  個が、またあるクラス  $C_j$  の特徴は別の  $n_j$  個が観測しているといった場合である. このとき各変量グループにもっとも相関の強い特徴量を抽出する変換ベクトルを見つけたし、それらのベクトルがつくる部分空間にパターンを射影するならばクラスを類別する重要な特徴が抽出されるであろう. しかし変量グループがクラスの特徴を主に観測しているとわかっているものの、どのようなグループになっているか事前には判明していないとしよう. このようなとき因子分析法における単純構造化の概念を導入し [定義 1] で示す基準量  $V$  を定め、これを最大にすることによって変量グループを明らかにする.

[定義 1]  $\eta$  の成分の平方  $\eta_j^2$  ( $j=1, \dots, n$ ) の分散

$$V = E[(\eta_j^2 - \mu)^2] = E[\eta_j^4] - \mu^2 \quad (32)$$

$$\text{ここに, } \mu = E[\eta_j^4]$$

を単純構造基準とよび、 $V$  の値を最大にする特徴量の構造を単純構造とよぶ.

$V$  が最大にされると  $\eta$  の成分の絶対値  $|\eta_j|$  が 1 に近いものと 0 に近いものとに分かれ、0.5 程度の値はなくなってしまふ. つまり特徴量の構造は一部の変量グループととくに相関が強くなり、他の変量グループとはとくに相関が弱くなり、中程度の相関がなくなってしまう. そして  $|\eta_j|$  の値が 1 に近い変量がグループとしてとりだされる. 定義より次の補題, 定理を得る.

[補題 1] パターン行列  $X$  より単純構造をなす特徴量ベクトル  $p_1 = X\xi_1$  を抽出するベクトル  $\xi_1$  は

$$\xi_1 = S^{-1}\zeta_1 \quad (33)$$

ここに,

$$\zeta_1 = \eta_1^{(3)} - \eta_1^T \eta_1 \eta_1 / n \quad (34)$$

$$\eta_1 = S^{-1} R S^{-1} \zeta_1 / \sqrt{\zeta_1^T S^{-1} R S^{-1} \zeta_1} \quad (35)$$

$$\eta_1 = (\eta_{11}, \eta_{21}, \dots, \eta_{n1})^T$$

$$\eta_1^{(3)} = (\eta_{11}^3, \eta_{21}^3, \dots, \eta_{n1}^3)^T$$

となる (証明は付録 [4]).

[補題 2]  $p_1$  と直交し、かつ単純構造をなす  $p_2 = X\xi_2$  を抽出するベクトル  $\xi_2$  は

$$\xi_2 = (I - \xi_1 \eta_1^T S / \sqrt{\xi_1^T R \xi_1}) S^{-1} \zeta_2 \quad (36)$$

ここに,

$$\zeta_2 = \eta_2^{(3)} - \eta_2^T \eta_2 \eta_2 / n \quad (37)$$

$$\eta_2 = S^{-1} R_1 S^{-1} \zeta_2 / \sqrt{\zeta_2^T S^{-1} R_1 S^{-1} \zeta_2} \quad (38)$$

$$\eta_2 = (\eta_{12}, \eta_{22}, \dots, \eta_{n2})^T$$

$$\eta_2^{(3)} = (\eta_{12}^3, \eta_{22}^3, \dots, \eta_{n2}^3)^T$$

となる (証明は付録 [4], [5] を参照).

一般に次の定理が成立する.

[定理 1]  $p_1, p_2, \dots, p_{i-1}$  と直交し、かつ単純構造をなす  $p_i = X\xi_i$  を抽出するベクトル  $\xi_i$  は

$$\xi_i = \left( I - \sum_{j=1}^{i-1} \xi_j \eta_j^T S / \sqrt{\xi_j^T R \xi_j} \right) S^{-1} \zeta_i \quad (39)$$

ここに,

$$\zeta_i = \eta_i^{(3)} - \eta_i^T \eta_i \eta_i / n \quad (40)$$

$$\eta_i = S^{-1} R_{i-1} S^{-1} \zeta_i / \sqrt{\zeta_i^T S^{-1} R_{i-1} S^{-1} \zeta_i} \quad (41)$$

$$\eta_i = (\eta_{1i}, \eta_{2i}, \dots, \eta_{ni})^T$$

$$\eta_i^{(3)} = (\eta_{1i}^3, \eta_{2i}^3, \dots, \eta_{ni}^3)^T$$

となる (証明は付録 [4], [5] を参照).

各  $\xi_i$  方向の軸は各クラスのもっとも重要な情報を担うことになる.  $\xi_i$  を正規化したものを  $\phi_i$  とし、 $\gamma$  値の大きい順に選び、選ばれた  $\phi_i$  がつくる部分空間にパターンを射影するのが単純構造変換である. 選ばれた  $l$  個の  $\phi$  を  $\phi_1, \phi_2, \dots, \phi_l$  とすれば

$$\Phi = (\phi_1, \phi_2, \dots, \phi_l)^T \quad (42)$$

はパターン  $x$  を

$$y = \Phi x \quad (43)$$

のように変換する. この  $\phi_i$  を単純化ベクトルとよぶことにする. 一方  $\xi_i$  を求めるには (39) 式からわかるように  $\zeta_i$  を求めねばならない.  $\zeta_i$  は (40) 式から求められるが右辺がまた  $\zeta_i$  の関数となっているので、最初  $\zeta_i$  は適当な初期値を代入し、新しく得られた  $\zeta_i$  を再び右辺の  $\zeta_i$  に代入するといった逐次反復計算を行い所要の精度まで収束させる必要がある. すなわち反復公式

$${}_{p+1}\zeta_i = {}_p\eta_i^{(3)} - {}_p\eta_i^T {}_p\eta_i {}_p\eta_i / n \quad (44)$$

$${}_p\eta_i = S^{-1} R_{i-1} S^{-1} {}_p\zeta_i / \sqrt{{}_p\zeta_i^T S^{-1} R_{i-1} S^{-1} {}_p\zeta_i} \quad (45)$$

をもちいて初期値  ${}_0\zeta_i$  から出発し  ${}_1\zeta_i, {}_2\zeta_i, \dots$  を計算し,

$$\max_j |{}_{p+1}\zeta_{ji} - {}_p\zeta_{ji}| < \epsilon$$

を満足するまで反復計算すればよい. なお  ${}_p\zeta_i = ({}_p\zeta_{1i}, {}_p\zeta_{2i}, \dots, {}_p\zeta_{ni})^T$  である.

### 4. K-L 変換との相違

3章で述べた単純構造変換と K-L 変換との相違に

ついて述べておく。K-L 変換は  $K$  個のパターンクラス  $C_i (i=1, \dots, K)$  が存在し、 $C_i$  の生起確率  $P_i$  を  $C_i$  のパターン自己相関行列に乘じた行列を

$$R^{(i)} = P_i E[\mathbf{x}_j^{(i)} \mathbf{x}_j^{(i)T}] \quad (i=1, \dots, K)$$

とすると、これらの行列を混合したものを

$$R = \sum_{i=1}^K R^{(i)}$$

とおき、この  $R$  の固有ベクトルがもちいられる。 $R$  は対称行列なので固有ベクトルどうしは互いに直交し、固有値は正または 0 である。固有値の大きい順に対応する固有ベクトルが選ばれ、選ばれた固有ベクトルがつくる部分空間にパターンを射影するのが K-L 変換である。各固有ベクトル方向の軸はすべての変量ともっとも相関の強い情報を、すなわち全パターンクラスともっとも相関の強い情報を担うことになる。K-L 変換は全パターンクラスにもっとも相関の強い空間を形成するといえよう。一方単純構造変換は各パターンクラスの特徴がそれぞれ別個の変量グループによって観測されているならば各パターンクラスにもっとも相関の強い特徴空間を形成した。したがって全パターンクラスの特徴が別個の変量グループによって観測されているときは K-L 変換よりも単純構造変換の方がクラス類別のための特徴をよく抽出する。

5. 計算機実験および検討

単純構造変換の有用性を確かめるため人工的なパターンをもちいて計算機実験を行った。表 1 は 4 次元の空間に 39 個のパターンが三つのクラスをなして分布している場合である。クラス  $C_1$  と  $C_2$  の特徴は主に変量 1 と変量 2 が、 $C_3$  は変量 3 と変量 4 が観測している。表 2 の構造ベクトル  $\eta_1$  と  $\eta_2$  は特徴量の構造が単純化されるのをよく示している。さらに表 3 の単純化ベクトル  $\phi_1$  はクラス  $C_1$  と  $C_2$  の、 $\phi_2$  は  $C_3$  の重要な特徴を抽出することをよく示している。表 4 は

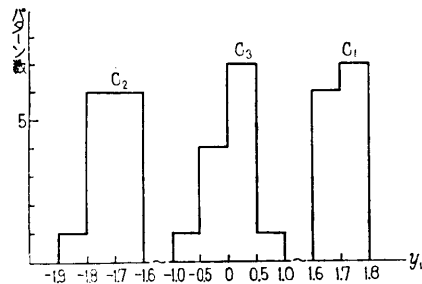


図 3 単純構造変換による特徴量  $y_1$  のヒストグラム  
Fig. 3 Histogram for feature  $y_1$  by simple-structure transformation.

表 1 4 次元パターン  
Table 1 4-dimensional patterns.

Class		$X_1$	$X_2$	$X_3$	$X_4$
$C_1$	1	1.227	1.211	-0.650	-0.666
	2	1.203	1.218	-1.050	-1.066
	3	1.155	1.168	-0.692	-0.674
	4	1.179	1.187	-0.251	-0.264
	5	1.203	1.206	-1.468	-1.432
	6	1.250	1.226	-1.050	-1.051
	7	1.203	1.233	-0.657	-0.645
	8	1.215	1.247	-0.271	-0.293
	9	1.234	1.238	0.097	0.110
	10	1.177	1.235	-1.040	-1.019
	11	1.186	1.206	-0.664	-0.670
	12	1.200	1.171	-0.299	-0.282
	13	1.203	1.159	-0.657	-0.670
$C_2$	14	-1.223	-1.122	-0.664	-0.681
	15	-1.211	-1.162	-1.040	-1.069
	16	-1.215	-1.181	-0.671	-0.670
	17	-1.199	-1.169	-0.320	-0.300
	18	-1.175	-1.189	-1.433	-1.396
	19	-1.187	-1.201	-1.016	-1.001
	20	-1.246	-1.210	-0.661	-0.672
	21	-1.251	-1.153	-0.282	-0.289
	22	-1.282	-1.165	0.080	0.046
	23	-1.270	-1.201	-1.019	-1.051
	24	-1.230	-1.210	-0.664	-0.659
	25	-1.175	-1.409	-0.317	-0.275
	26	-1.170	-1.363	-0.657	-0.681
	$C_3$	27	0.021	0.008	1.279
28		-0.231	-0.248	1.422	1.121
29		0.009	-0.019	1.488	1.157
30		0.283	0.292	1.384	1.390
31		-0.486	-0.483	1.419	1.455
32		-0.241	-0.251	1.505	1.516
33		0.019	0.008	1.568	1.613
34		0.261	0.245	1.422	1.624
35		0.506	0.491	1.081	1.113
36		-0.248	-0.277	1.384	1.398
37		0.016	-0.002	1.057	1.139
38		0.273	0.269	1.213	1.394
39		0.019	-0.004	1.095	1.171

表 2 構造ベクトル  
Table 2 Structure vectors.

	$\eta_1$	$\eta_2$
1	0.999	0.001
2	0.999	-0.006
3	0.001	0.999
4	0.007	0.998
$r$	1.998	1.996

$\phi_1, \phi_2$  によって変換されたパターン  $y_i = (\phi_1, \phi_2)^T x_i$  を示している。 $\phi_1$  による特徴量  $y_1$  をみると  $C_1$  と  $C_2$  の特徴がよく抽出され、結果として  $C_3$  ともよく類別されている。 $\phi_2$  による  $y_2$  をみると  $C_3$  の特徴がよく抽出され、 $C_1, C_2$  のそれとよく類別されている。図 3

表 3 単純化ベクトル  
Table 3 Simplistic vectors.

	$\phi_1$	$\phi_2$
1	0.703	-0.005
2	0.710	-0.000
3	0.003	0.713
4	-0.000	0.700

表 4 単純構造変換による特徴量  
Table 4 Features by simple-structure transformation.

Class		$v_1$	$v_2$
$C_1$	1	1.721	-0.937
	2	1.708	-1.503
	3	1.640	-0.972
	4	1.672	-0.371
	5	1.698	-2.057
	6	1.747	-1.492
	7	1.720	-0.927
	8	1.740	-0.405
	9	1.748	0.138
	10	1.702	-1.462
	11	1.689	-0.950
	12	1.675	-0.418
	13	1.667	-0.945
$C_2$	14	-1.659	-0.943
	15	-1.681	-1.483
	16	-1.696	-0.940
	17	-1.675	-0.431
	18	-1.676	-1.993
	19	-1.691	-1.419
	20	-1.738	-0.935
	21	-1.700	-0.396
	22	-1.729	0.096
	23	-1.750	-1.455
	24	-1.727	-0.928
	25	-1.829	-0.411
	26	-1.793	-0.938
	$C_3$	27	0.024
28		-0.333	1.801
29		-0.002	1.872
30		0.411	1.959
31		-0.680	2.034
32		-0.343	2.137
33		0.024	2.248
34		0.362	2.150
35		0.708	1.548
36		-0.366	1.968
37		0.013	1.552
38		0.387	1.840
39		0.014	1.601

表 5 正規化された固有ベクトル  
Table 5 Normalied eigen vectors.

	$\phi_1$	$\phi_2$
1	0.568	-0.419
2	0.565	-0.424
3	0.420	0.568
4	0.424	0.565
固有値	2.001	1.993

表 6 K-L 変換による特徴量  
Table 6 Feature by Karhunen-Loève transform.

Class		$v_1$	$v_2$
$C_1$	1	0.827	-1.775
	2	0.479	-2.222
	3	0.740	-1.755
	4	1.124	-1.291
	5	0.141	-2.662
	6	0.517	-2.237
	7	0.831	-1.767
	8	1.158	-1.359
	9	1.489	-0.926
	10	0.498	-2.186
	11	0.793	-1.766
	12	1.009	-1.330
	13	0.779	-1.749
$C_2$	14	-1.898	0.227
	15	-2.236	-0.194
	16	-1.925	0.251
	17	-1.604	0.648
	18	-2.535	-0.606
	19	-2.206	-0.135
	20	-1.955	0.281
	21	-1.604	0.691
	22	-1.334	1.104
	23	-2.275	-0.130
	24	-1.942	0.279
	25	-1.715	0.755
	26	-2.001	0.311
	$C_3$	27	1.075
28		0.801	1.645
29		1.110	1.504
30		1.497	1.330
31		0.664	2.039
32		0.996	1.921
33		1.358	1.792
34		1.573	1.513
35		1.491	0.823
36		0.877	1.799
37		0.935	1.259
38		1.408	1.249
39		0.965	1.278

は  $v_1$  に対するヒストグラムで三つのクラスがよく類別されている有様を明示したものである。

次に同じデータをもちい、パターンクラス生起確率を  $P_1=P_2=P_3=1/3$  にとって K-L 変換を行ってみた。表 5 の  $\phi_1, \phi_2$  は得られた固有ベクトルを正規化し

たものである。 $\phi_1, \phi_2$  の方向に抽出される特徴量が全パターンクラスに相関の強いものとなることがよくわかる。表 6 は  $\phi_1, \phi_2$  による特徴量である。 $\phi_1$  による  $v_1$  をみると  $C_2$  は他の  $C_1, C_3$  とよく類別されているが、 $C_1$  と  $C_3$  はひどく重なりがみられる。 $\phi_2$  による

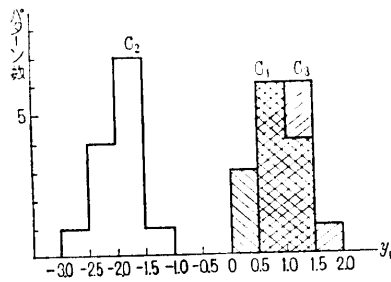


図4 K-L変換による特徴量  $y_1$  のヒストグラム  
Fig. 4 Histogram for feature  $y_1$  by Karhunen-Loève transform.

$y_2$  をみると  $C_1$  が他の  $C_2, C_3$  とよく類別されているが,  $C_2, C_3$  は重なりがみられる. 図4は  $y_1$  に対するヒストグラムである.

この例では三つのクラスを類別するのに単純構造変換では1次元部分空間にパターンを射影するだけでよいが, K-L変換では少なくとも2次元の部分空間に射影する必要がある. この実験結果からもわかるように単純構造変換が形成する部分空間の各軸は少なくとも一つ(例では  $\phi_1$  方向の軸は二つ)のクラスのもっとも重要な情報を担うことになる. それに対し K-L変換での軸は全クラスの情報を担っている. K-L変換はクラス類別のための特徴を抽出するのに必ずしも最適ではないことがわかる. 一般に  $K$  個のパターンクラスが分布し, 各クラスの特徴がそれぞれ別個の変量グループによって観測されている場合, 高々  $K$  個の単純化ベクトルがつくる部分空間にパターンを射影するのがもっとも適当といえる.

## 6. むすび

パターンクラスの特徴が主に一部の変量グループによって観測されているとわかっているものの, どのようなグループになっているのか不明のとき, 単純構造基準を最大にすることによってグループを見つけだし, 同時にその変量グループにもっとも相関の強い特徴量を抽出する単純化ベクトルを見いだす方法について述べた. そしてこれらのベクトルがつくる部分空間にパターンを射影するならばクラスを類別する重要な特徴が得られることを実験により確かめ, 従来の K-L変換よりも有用であることを示した. 本方法のもっとも大きな特徴は単純化ベクトルがつくる部分空間の各軸は少なくとも一つのクラスの重要な情報を担うという点であろう. 一般に  $K$  個のクラスならば高々  $K$  個の単純化ベクトルでよい. 本方法は高次元多クラス問

題においてクラスの特徴が主として一部の変量グループによって観測されている場合の一つの特徴抽出法として有効に働くものと思われる. 今後実在する高次元パターンに対し本方法の有用性を確かめていきたい.

最後に日ごろより有益な討論をしていただく当研究室の金岡泰保氏に深謝する.

## 参考文献

- 1) Watanabe, S.: Karhunen-Loève Expansion and Factor Analysis, Trans. 4th Prague Conf. on Information Theory, pp. 635-660 (1965).
- 2) Chien, Y. T. and Fu, K. F.: On the Generalized Karhunen-Loève Expansion, IEEE Trans. Inform. Theory, IT-13, pp. 518-520 (1967).
- 3) Fukunaga, K. and Koontz, W. L. G.: Application of the Karhunen-Loève Expansion to Selection and Ordering, IEEE Trans. Comput., C-19, pp. 311-318 (1970).
- 4) Foley, D. H. and Sammon, J. W. Jr.: An Optimal Set of Discriminant Vectors, IEEE Trans. Comput., C-24, pp. 281-289 (1975).
- 5) 芝 祐順: 因子分析法, p. 424, 東京大学出版会 (1975).
- 6) 遠山 啓: 行列論, p. 266, 共立出版 (1975).
- 7) 富田, 野口, 大泉: Karhunen-Loève 直交系による教師なしの類別, 電子通信学会論文誌 C, Vol. 54-C, No. 8, pp. 767-774 (1971).

## 付 録

[1] 一般にパターン認識系においては観測, 特徴抽出, 識別の各処理系がシステムとして互に関連を保ち, 一つの処理系を構成する. しかし, 対象とするパターン集合について事前に何らかの知識を与えられていない場合は, いわゆる教師なしの構成となるため, その処理機構は非常に複雑で, かつ実用的でない<sup>7)</sup>. しかし, 対象のパターン集合があらかじめ何らかの知識, 例えば  $n$ 次元空間に分布を構成しているとか, パターンの塊がいくつあるとか等が与えられていれば, その知識を基にパターン処理系を構成する設計方法を組み立て, さらに一般的な場合へと拡張するのが妥当である. 本論文はこのような立場から与えられたパターン集合はあらかじめ何らかの知識, すなわち  $n$ 次元空間にパターンの塊(これをパターンクラス, あるいは単にクラスとよぶ)が  $K$  個存在すると仮定し, これらの塊がパターン集合全体を特徴づけている, すなわち塊の一つの特徴とみなして, それを抽出することを考察する.

ところで各クラスの相違は各クラスに属するパター



ンの観測値, すなわちパターンベクトルを構成する  $n$  個の変量の観測値に起因するのは明らかである. しかし,  $n$  個の変量のすべてが同程度によくクラス間の相違を観測しているとはかぎらない. 各変量の観測値はクラス間の相違に異なった寄与を与えると考えるのが自然であろう. 換言すれば各クラスにおいては他クラスとの相違をよく観測している変量とそうでない変量が存在する. 本論文ではこのような観点から各クラスにおける他クラスとの相違特徴, すなわちクラス特有の特徴をよく観測している変量を因子分析法における単純構造の概念を導入して見つけ, それらの変量にもっとも相関の強い特徴を抽出する方法について述べている.

[2] いま  $\mathbf{p}_2^*$  を  $\mathbf{X}$  から任意ベクトル  $\xi_2^*$  をもちいて得られるベクトルとする. すなわち (4) 式の  $\mathbf{p}_1$  と同様に

$$\mathbf{p}_2^* = \mathbf{X}\xi_2^*. \quad (\text{A1})$$

ただし,  $\mathbf{p}_2^*$  と  $\mathbf{p}_1$  の間には  $\mathbf{p}_2^* = c\mathbf{p}_1$  を満たす定数  $c$  は  $c=0$  のみとする.  $\mathbf{p}_2$  を

$$\mathbf{p}_2 = \mathbf{p}_2^* - c\mathbf{p}_1 \quad (\text{A2})$$

とおき,  $\mathbf{p}_1^T \mathbf{p}_2 = 0$  を満足するように  $c$  を定めると

$$c = \mathbf{p}_1^T \mathbf{p}_2^* / \mathbf{p}_1^T \mathbf{p}_1 = \xi_1^T \mathbf{R}\xi_2^* / \xi_1^T \mathbf{R}\xi_1 \quad (\text{A3})$$

となる. したがって

$$\begin{aligned} \mathbf{p}_2 &= \mathbf{X}\xi_2^* - (\xi_1^T \mathbf{R}\xi_2^* / \xi_1^T \mathbf{R}\xi_1) \mathbf{X}\xi_1 \\ &= \mathbf{X}(\mathbf{I} - \xi_1 \xi_1^T \mathbf{R} / \xi_1^T \mathbf{R}\xi_1) \xi_2^* \\ &= \mathbf{X}(\mathbf{I} - \xi_1 \eta_1^T \mathbf{S} / \sqrt{\xi_1^T \mathbf{R}\xi_1}) \xi_2^* \\ &= \mathbf{X}_1 \xi_2^* \end{aligned}$$

となる (証明終).

[3]  $\mathbf{p}_1$  および  $\mathbf{p}_2^*$  に準じて

$$\begin{aligned} \mathbf{p}_{i-1} &= \mathbf{X}\xi_{i-1} = \mathbf{X}_{i-2} \xi_{i-1}^* \\ \mathbf{p}_i^* &= \mathbf{X}_{i-2} \xi_i^* \end{aligned}$$

とおく. ただし,  $\mathbf{p}_i^* = c\mathbf{p}_{i-1}$  が成立する定数  $c$  は  $c=0$  だけとする.

$$\mathbf{p}_i = \mathbf{p}_i^* - c\mathbf{p}_{i-1}$$

とおき,  $\mathbf{p}_{i-1}^T \mathbf{p}_i = 0$  を満足するように  $c$  を定めると

$$\begin{aligned} c &= \mathbf{p}_{i-1}^T \mathbf{p}_i^* / \mathbf{p}_{i-1}^T \mathbf{p}_{i-1} \\ &= \xi_{i-1}^{*T} \mathbf{R}_{i-2} \xi_i^* / \xi_{i-1}^{*T} \mathbf{R}_{i-2} \xi_{i-1}^* \end{aligned}$$

となる. したがって

$$\begin{aligned} \mathbf{p}_i &= \mathbf{p}_i^* - (\xi_{i-1}^{*T} \mathbf{R}_{i-2} \xi_i^* / \xi_{i-1}^{*T} \mathbf{R}_{i-2} \xi_{i-1}^*) \mathbf{p}_{i-1} \\ &= (\mathbf{X}_{i-2} \\ &\quad - \mathbf{X}_{i-2} \xi_{i-1}^* \xi_{i-1}^{*T} \mathbf{R}_{i-2} / \xi_{i-1}^{*T} \mathbf{R}_{i-2} \xi_{i-1}^*) \xi_i^* \\ &= (\mathbf{X}_{i-2} - \mathbf{X}\xi_{i-1} \eta_{i-1}^T \mathbf{S} / \sqrt{\xi_{i-1}^T \mathbf{R}\xi_{i-1}}) \xi_i^* \end{aligned}$$

$\mathbf{X}_{i-2}$  に (20) 式の  $\mathbf{X}_{i-2}$  の場合を代入すれば

$$\mathbf{p}_i = \mathbf{X} \left( \mathbf{I} - \sum_{j=1}^{i-1} \xi_j \eta_j^T \mathbf{S} / \sqrt{\xi_j^T \mathbf{R}\xi_j} \right) \xi_i^*$$

となる (証明終).

[4] 以下の証明の中で式を簡潔にするために (32) 式を  $n$  倍したものを

$$W = nV = nE[\eta_j^4] - n(E[\eta_j^2])^2 \quad (\text{B1})$$

と表わすことにする.  $\mathbf{p}_1$  を単純構造にするため  $W$  を

$$W_1 = \sum_{j=1}^n \eta_{j1}^4 - \left( \sum_{j=1}^n \eta_{j1}^2 \right)^2 / n \quad (\text{B2})$$

とおき, これを最大にする  $\eta_1$  を  $\mathbf{R}$  の成分  $r_{jk}$  が

$$r_{jk} = s_{js} s_{sk} \sum_{p=1}^m \eta_{jp} \eta_{kp} \quad (\text{B3})$$

となる条件のもとで求めてみる. そのために

$$4U = W_1 - 2 \sum_{j=1}^n \sum_{k=1}^n \mu_{jk} \left( \sum_{p=1}^m \eta_{jp} \eta_{kp} - r_{jk} / s_{js} s_{sk} \right) \quad (\text{B4})$$

のような値  $U$  を考え,  $U$  を各  $\eta_{jp}$  ( $p=1, \dots, m$ ) について偏微分し, 結果をすべて 0 とおく. ここで,  $\mu_{jk}$  はラグランジュの乗数である. まず  $\partial U / \partial \eta_{j1} = 0$  を実行すると

$$\eta_{j1}^3 - \eta_{j1} \sum_{k=1}^n \eta_{k1}^2 / n = \sum_{k=1}^n \mu_{jk} \eta_{k1} \left( = \sum_{k=1}^n \mu_{kj} \eta_{k1} \right). \quad (\text{B5})$$

$\partial U / \partial \eta_{jp} = 0$  ( $p \neq 1$ ) を実行すると

$$\sum_{k=1}^n \mu_{jk} \eta_{kp} = 0 \quad (\text{B6})$$

を得る. (B5) と (B6) を一つの式にまとめると

$$\delta_{1p} \left( \eta_{jp}^3 - \eta_{jp} \sum_{k=1}^n \eta_{k1}^2 / n \right) = \sum_{k=1}^n \mu_{jk} \eta_{kp} \quad (\text{B7})$$

ここに,  $\delta_{1p}$  は  $p=1$  のとき 1,  $p \neq 1$  のとき 0 の値をとる. 両辺に  $\eta_{j1}$  を乗じて  $j$  について和をとると

$$\begin{aligned} \delta_{1p} \left( \sum_{j=1}^n \eta_{j1}^4 - \sum_{j=1}^n \eta_{j1}^2 \sum_{k=1}^n \eta_{k1}^2 / n \right) \\ = \sum_{j=1}^n \sum_{k=1}^n \mu_{jk} \eta_{j1} \eta_{kp} \end{aligned}$$

となり, これは

$$\delta_{1p} W_1 = \sum_{k=1}^n \eta_{kp} \sum_{j=1}^n \mu_{jk} \eta_{j1}$$

と書ける. 右辺の  $\sum_{j=1}^n \mu_{jk} \eta_{j1}$  は (B5) の  $j$  と  $k$  を交換したものにほかならない. したがって

$$\delta_{1p} W_1 = \sum_{k=1}^n \eta_{kp} \zeta_{k1}.$$

ここに,  $\zeta_{k1} = \eta_{k1}^3 - \eta_{k1} \sum_{k=1}^n \eta_{k1}^2 / n$  である.

さらに、両辺に  $\eta_{1p}$  を乗じて  $p$  について和をとると

$$\begin{aligned} \sum_{p=1}^m \delta_{1p} \eta_{1p} W_1 &= \sum_{k=1}^n \zeta_{k1} \sum_{p=1}^m \eta_{1p} \eta_{kp} \\ &= \sum_{k=1}^n \zeta_{k1} r_{1k} / s_{1s_k} \end{aligned}$$

となる。したがって

$$\eta_{11} W_1 = \sum_{k=1}^n \zeta_{k1} r_{1k} / s_{1s_k}$$

となり、まとめて行列とベクトルで表わすと

$$W_1 \eta_1 = S^{-1} R S^{-1} \zeta_1 \tag{B8}$$

ここに、 $\zeta_1 = \eta_1^{(3)} - \eta_1^T \eta_1 \eta_1 / n$ ,

$$\eta_1^{(3)} = (\eta_{11}^3, \eta_{21}^3, \dots, \eta_{m1}^3)^T.$$

一方、(B8)の両辺に  $\zeta_1^T$  を乗じると

$$W_1 \zeta_1^T \eta_1 = \zeta_1^T S^{-1} R S^{-1} \zeta_1.$$

また、 $\zeta_1^T \eta_1 = W_1$  より

$$W_1^2 = \zeta_1^T S^{-1} R S^{-1} \zeta_1.$$

したがって

$$W_1 = \sqrt{\zeta_1^T S^{-1} R S^{-1} \zeta_1}.$$

結局(B8)は

$$\eta_1 = S^{-1} R S^{-1} \zeta_1 / \sqrt{\zeta_1^T S^{-1} R S^{-1} \zeta_1}.$$

これと(7)式  $\eta_1 = S^{-1} R \xi_1 / \sqrt{\xi_1^T R \xi_1}$  を比較して

$$\xi_1 = S^{-1} \zeta_1$$

を得る (証明終).

[5]  $p_1$  と直交し、かつ単純構造をなす  $p_2$  を得るために

$$W_2 = \sum_{j=1}^n \eta_{j2}^4 - \left( \sum_{j=1}^n \eta_{j2}^2 \right)^2 / n \tag{C1}$$

とおき、これを最大にする  $\eta_2$  を  $R_1$  の成分  $r_{jk}$  が

$$r_{jk} = r_{jk} - s_{js_k} \eta_{j1} \eta_{k1} = s_{js_k} \sum_{p=2}^m \eta_{jp} \eta_{kp} \tag{C2}$$

となる条件のもとで求めればよい。(B4)に対応する

$$4U = W_2 - 2 \sum_{j=1}^n \sum_{k=1}^n \mu_{jk} \left( \sum_{p=2}^m \eta_{jp} \eta_{kp} - r_{jk} / s_{js_k} \right) \tag{C3}$$

なる値  $U$  を考え、これを各  $\eta_{jp}$  ( $p=2, \dots, m$ ) について偏微分し、結果をすべて0とおく。以下前と同様の手法で進めれば

$$\eta_2 = S^{-1} R_1 S^{-1} \zeta_2 / \sqrt{\zeta_2^T S^{-1} R_1 S^{-1} \zeta_2}$$

を得る。これと(15)式  $\eta_2 = S^{-1} R_1 \xi_2^* / \sqrt{\xi_2^{*T} R_1 \xi_2^*}$  を比較して、 $\xi_2^* = S^{-1} \zeta_2$  が得られ、(12)式より

$$\xi_2 = (I - \xi_1 \eta_1^T S / \sqrt{\xi_1^T R \xi_1}) S^{-1} \zeta_2$$

を得る (証明終)。以下同様にして証明でき、一般に  $p_1, p_2, \dots, p_{i-1}$  と直交し、かつ単純構造をなす  $p_i$  を抽出する  $\xi_i$  を順次得る。

(昭和53年8月25日受付)

(昭和54年5月17日採録)