



| | |
|------------------|--|
| Title | パーソナルコンピュータによる多次元データの動的表示法 |
| Author(s) | 水田, 正弘; 佐藤, 義治; 伊達, 惇; 河口, 至商 |
| Citation | 北海道大學工學部研究報告 = Bulletin of the Faculty of Engineering, Hokkaido University, 109: 55-64 |
| Issue Date | 1982-07-30 |
| Doc URL | http://hdl.handle.net/2115/41733 |
| Type | bulletin (article) |
| File Information | 109_55-64.pdf |



[Instructions for use](#)

パーソナルコンピュータによる多次元
データの動的表示法

水田正弘* 佐藤義治*
伊達 惇** 河口至商*
(昭和57年3月31日受理)

**Dynamic Representation of Multi-Dimensional Data
Using a Personal Computer**

Masahiro MIZUTA, Yoshiharu SATO, Tsutomu DATE
and Michiaki KAWAGUCHI
(Received March 31, 1982)

Abstract

The representation of data by a 2-dimensional plane often makes it possible to understand the data structure intuitively. Several graphical methods have been studied to this end.

We will present one of the methods that represent multi-dimensional data as the motion of points. A personal computer is available for this method. The basic concept of the method is that the plane is moved continuously from one direction to another in a p -dimensional vector space: the pair of independent vectors (X, Y) is altered from one direction to another, and we use the motion of the points that are projections of the data on the moving plane. This motion of the plane is designated by the user. The method is interactive and assists in heuristic investigations in the data structure. Also some concrete examples are given.

1. 序 論

多次元データを平面上に表現することにより、データの構造を直観的に把握できる場合が多い。このようにデータをグラフ化して解析する方法は『多変量グラフ解析法』と呼ばれており、Sammonの非線型変換プロット¹⁾、Chernoffの顔グラフ²⁾、脇本・田栗の星座グラフ³⁾、Andrewsの三角多項式グラフ⁴⁾、新保・宮腰の標本化関数グラフ⁵⁾などがある。本研究では、パーソナルコンピュータの特性を利用した多次元データの動的表示法を提案する。この手法は会話形式で、利用者が指定した平面に多次元データが射影されるよう射影する平面を連続的に移動させ、データを画面上の点の動きとして表現するものである。

Andrewsの指摘⁶⁾などにもあるように、パーソナルコンピュータは表示速度、会話処理の容易さ

* 情報工学専攻 情報数理工学第一講座

** 情報工学専攻 情報科学

などから、データのグラフ解析に適していると言える。本手法はパーソナルコンピュータを利用することを想定し、複雑な図形を用いずアルゴリズムも容易なものとしている。

本手法がデータ構造の解析に有効であること、またデータ解析における種々な手法の比較に利用できる可能性を持つことを具体的な例を用いて示す。

2. 手法の基本的説明

以下で取り扱う多次元データを N 個の p 次元ベクトル: $\{x_a\}$ ($a=1, 2, \dots, N$) と書く。またデータを表示する装置としては (カラー) グラフィックディスプレイを用い、その画面 (以下、CRT と呼ぶ) の座標を (x, y) と表す。2つの独立な p 次元ベクトル X と Y によってデータ $\{x_a\}$ を射影するとは、CRT 上の座標 $(X \cdot x_a, Y \cdot x_a)$ ($a=1, 2, \dots, N$) に点を表示することである (ここで、 $X \cdot x_a$ 等は内積を意味する)。

本論で提案する手法はデータの射影される平面を連続的に移動させることにより、多次元データを CRT 上の点の動きとして表現するものである。 X_1 と Y_1 は現在表示している平面を表すベクトルとし、 X_2 と Y_2 は連続的に平面を移動させたときの最終的な表示平面を表すベクトルとする (図 1)。

結果の解析が容易になるように、次の仮定を設ける。

$$\begin{cases} X_i \cdot Y_i = 0 \\ |X_i| = |Y_i| = 1 \end{cases} \quad (i = 1, 2)$$

多次元データを動的に表示するためには X_1 を X_2 に、 Y_1 を Y_2 にそれぞれ連続的に移すベクトル値関数 $X^*(t), Y^*(t)$ が必要である。この $X^*(t)$ と $Y^*(t)$ は少なくとも次の条件を満たすべきである。

条件 1 $X^*(t)$ と $Y^*(t)$ は t に関して連続なベクトル値関数である ($0 \leq t \leq 1$)。

条件 2 $X^*(0) = X_1, X^*(1) = X_2,$
 $Y^*(0) = Y_1, Y^*(1) = Y_2.$

条件 3 $X^*(t) \perp Y^*(t) \quad (0 \leq t \leq 1).$

条件 4 $|X^*(t)| = |Y^*(t)| = 1 \quad (0 \leq t \leq 1).$

これらの条件を満たす $X^*(t), Y^*(t)$ を次章で述べる。

3. $X^*(t), Y^*(t)$ の定義

$X^*(t)$ および $Y^*(t)$ を定義するための補題を準備する。

補題 互いに独立でノルムの等しい p 次元ベクトル V_1, V_2 について曲線

$$V(t) = \frac{\sin(\theta - \theta t)}{\sin \theta} V_1 + \frac{\sin(\theta t)}{\sin \theta} V_2 \quad (0 \leq t \leq 1)$$

は、 V_1 と V_2 によって張られる平面の上に存在し、 V_1 および V_2 を端点とする円弧となる (図 2)。ただし、 θ は V_1 と V_2 とのなす角とする ($0 < \theta \leq \pi/2$)。

(証明) $V(t)$ が V_1 と V_2 によって張られる平面上にあることは明らかである。また、

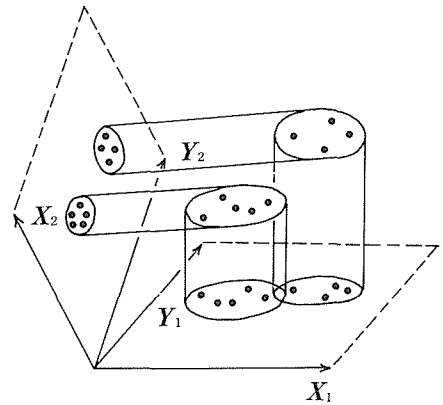


図 1 データの 2 平面への射影

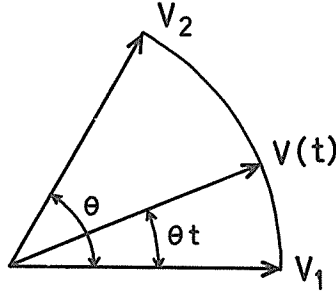


図2 ベクトル間を結ぶ円弧

$$\begin{aligned}
 |V(t)|^2 &= \left(\frac{\sin(\theta - \theta t)}{\sin\theta} V_1 + \frac{\sin(\theta t)}{\sin\theta} V_2 \right) \cdot \left(\frac{\sin(\theta - \theta t)}{\sin\theta} V_1 + \frac{\sin(\theta t)}{\sin\theta} V_2 \right) \\
 &= \frac{\sin^2(\theta - \theta t)}{\sin^2\theta} |V_1|^2 + \frac{2\sin(\theta - \theta t) \cdot \sin(\theta t)}{\sin^2\theta} V_1 \cdot V_2 + \frac{\sin^2(\theta t)}{\sin^2\theta} |V_2|^2 \\
 &= \frac{\sin^2(\theta - \theta t) + 2\sin(\theta - \theta t) \cdot \sin(\theta t) \cdot \cos\theta + \sin^2(\theta t)}{\sin^2\theta} |V_1|^2 \\
 &= |V_1|^2
 \end{aligned}$$

および, $V(0) = V_1$, $V(1) = V_2$ より補題の後半も成立する。

Q. E. D.

上記の補題より $V(t)$ は V_1 と V_2 を結ぶ自然な曲線であることが分る。この補題を利用することにより, $X_1 + Y_1$ と $X_2 + Y_2$ を結ぶ $W(t)$ および $X_1 - Y_1$ と $X_2 - Y_2$ を結ぶ $S(t)$ を次のように定義することができる。

$$W(t) = \frac{\sin(\theta_w - \theta_w t)}{\sin\theta_w} (X_1 + Y_1) + \frac{\sin(\theta_w t)}{\sin\theta_w} (X_2 + Y_2) \quad (1)$$

$$S(t) = \frac{\sin(\theta_s - \theta_s t)}{\sin\theta_s} (X_1 - Y_1) + \frac{\sin(\theta_s t)}{\sin\theta_s} (X_2 - Y_2) \quad (2)$$

ただし,

$$\cos\theta_w = (X_1 + Y_1) \cdot (X_2 + Y_2) / 2, \quad 0 \leq \theta_w \leq \pi,$$

$$\cos\theta_s = (X_1 - Y_1) \cdot (X_2 - Y_2) / 2, \quad 0 \leq \theta_s \leq \pi,$$

とする。ここで以下の式を満たす $X(t)$, $Y(t)$ を考える。

$$W(t) = X(t) + Y(t) \quad (3)$$

$$S(t) = X(t) - Y(t) \quad (4)$$

(1), (2), (3), (4)式より,

$$\begin{aligned}
 X(t) &= \frac{1}{2} \left\{ \left(\frac{\sin(\theta_w - \theta_w t)}{\sin\theta_w} + \frac{\sin(\theta_s - \theta_s t)}{\sin\theta_s} \right) X_1 + \left(\frac{\sin(\theta_w t)}{\sin\theta_w} + \frac{\sin(\theta_s t)}{\sin\theta_s} \right) X_2 \right. \\
 &\quad \left. + \left(\frac{\sin(\theta_w - \theta_w t)}{\sin\theta_w} - \frac{\sin(\theta_s - \theta_s t)}{\sin\theta_s} \right) Y_1 + \left(\frac{\sin(\theta_w t)}{\sin\theta_w} - \frac{\sin(\theta_s t)}{\sin\theta_s} \right) Y_2 \right\} \quad (5)
 \end{aligned}$$

$$\begin{aligned}
 Y(t) &= \frac{1}{2} \left\{ \left(\frac{\sin(\theta_w - \theta_w t)}{\sin\theta_w} - \frac{\sin(\theta_s - \theta_s t)}{\sin\theta_s} \right) X_1 + \left(\frac{\sin(\theta_w t)}{\sin\theta_w} - \frac{\sin(\theta_s t)}{\sin\theta_s} \right) X_2 \right. \\
 &\quad \left. + \left(\frac{\sin(\theta_w - \theta_w t)}{\sin\theta_w} + \frac{\sin(\theta_s - \theta_s t)}{\sin\theta_s} \right) Y_1 + \left(\frac{\sin(\theta_w t)}{\sin\theta_w} + \frac{\sin(\theta_s t)}{\sin\theta_s} \right) Y_2 \right\} \quad (6)
 \end{aligned}$$

となる。この $X(t)$ と $Y(t)$ を正規化したベクトル値関数を $X^*(t)$, $Y^*(t)$ とする。すなわち,

$$X^*(t) = \frac{X(t)}{|X(t)|} \quad (7)$$

$$Y^*(t) = \frac{Y(t)}{|X(t)|} \quad (8)$$

以上の定義は $0 < \theta_w \leq \pi/2$ かつ $0 < \theta_s \leq \pi/2$ の場合のみ考えたが、これ以外の場合には以下のようにして定義する。

$\theta_w = 0$ または $\theta_s = 0$ の場合には極限を考え、(5)と(6)式において、

$$\frac{\sin(\theta_w t)}{\sin \theta_w} = t, \quad \frac{\sin(\theta_s - \theta_s t)}{\sin \theta_s} = 1 - t \quad (9)$$

などとみなす。

$\pi/2 < \theta_w \leq \pi$ の場合には、

$$\begin{cases} Y_2 := -X_2 \\ X_2 := -Y_2 \end{cases}$$

と変換して $0 \leq \theta_w < \pi/2$ とする。また $\pi/2 < \theta_s \leq \pi$ の場合には X_2 と Y_2 を交換して $\theta_s = 0$ とする。 $X_1 = X_2$ かつ $Y_1 = -Y_2$ の場合には、 $Y_2 := -Y_2, X_1 = -X_2$ かつ $Y_1 = Y_2$ の場合には $X_2 := -X_2$ とする。これらの変換は結果の解析に対して影響を与えない自然なものである。

(7), (8)によって定義された $X^*(t)$ と $Y^*(t)$ が前述の条件 1, 2, 4 を満たすことを示すのは容易である(ただし、条件 2 における X_2 と Y_2 は上述の変換後のものとする)。また次の定理より、条件 3 を満たすことが示される。

定理 1 (7), (8)により定義された $X^*(t)$ と $Y^*(t)$ は $0 \leq t \leq 1$ において、

$$X^*(t) \cdot Y^*(t) = 0$$

となる。

(証明) 補題および(1), (2), (3), (4)式より、

$$|X(t) + Y(t)|^2 = |X_1 + Y_1|^2 = |X_2 + Y_2|^2 = 2$$

$$|X(t) - Y(t)|^2 = |X_1 - Y_1|^2 = |X_2 - Y_2|^2 = 2$$

が示される。従って、両辺を展開することにより、

$$|X(t)|^2 + 2X(t) \cdot Y(t) + |Y(t)|^2 = 2$$

$$|X(t)|^2 - 2X(t) \cdot Y(t) + |Y(t)|^2 = 2$$

が成立する。辺々を引くことにより、

$$X(t) \cdot Y(t) = 0$$

ゆえに

$$X^*(t) \cdot Y^*(t) = 0$$

Q. E. D.

注意: 上記の証明より

$$|X(t)|^2 + |Y(t)|^2 = 2$$

また、 $\theta_w \neq 0$ かつ $\theta_s \neq 0$ のとき、

$$\begin{aligned} |X(t)|^2 &= \frac{1}{4} |W(t) + S(t)|^2 \\ &= \frac{1}{4} \left\{ |W(t)|^2 + |S(t)|^2 + 2W(t) \cdot S(t) \right\} \\ &= \frac{1}{4} \left\{ 4 + 2W(t) \cdot S(t) \right\} \end{aligned}$$

$$= 1 + \frac{1}{2} \left\{ \frac{\sin(\theta_w - \theta_w t) \cdot \sin(\theta_s t)}{\sin \theta_w \cdot \sin \theta_s} (X_1 + Y_1) \cdot (X_2 - Y_2) \right. \\ \left. + \frac{\sin(\theta_w t) \cdot \sin(\theta_s - \theta_s t)}{\sin \theta_w \cdot \sin \theta_s} (X_1 - Y_1) \cdot (X_2 + Y_2) \right\}$$

ここで、 $0 < \theta < \pi/2$, $0 \leq t \leq 1$ なる任意の θ , t に対して、

$$\frac{\sin(\theta t)}{\sin \theta} \leq \sin\left(\frac{\pi}{2} t\right)$$

が成立する (ただし、等号が成立するのは $t=0, 1$ のときに限る) ので、

$$|X(t)|^2 \leq 1 + \frac{1}{2} \sin\left(\frac{\pi}{2} t\right) \cdot \sin\left(\frac{\pi}{2} (1-t)\right) \left\{ |(X_1 + Y_1) \cdot (X_2 - Y_2)| + |(X_1 - Y_1) \cdot (X_2 + Y_2)| \right\} \\ = 1 + \frac{1}{2} \sin\left(\frac{\pi}{2} t\right) \cdot \cos\left(\frac{\pi}{2} t\right) \left\{ |(X_1 + Y_1) \cdot (X_2 - Y_2)| + |(X_1 - Y_1) \cdot (X_2 + Y_2)| \right\} \\ \leq 1 + \left\{ |(X_1 + Y_1) \cdot (X_2 - Y_2)| + |(X_1 - Y_1) \cdot (X_2 + Y_2)| \right\} / 4$$

となる。ここで最初の不等号が等号で成立するのは $t=0, 1$ のみ、次の不等号が等号で成立するのは $t=1/2$ の場合に限るので、

$$|X(t)|^2 < 1 + \left\{ |(X_1 + Y_1) \cdot (X_2 - Y_2)| + |(X_1 - Y_1) \cdot (X_2 + Y_2)| \right\} / 4 \\ \leq 2$$

と導くことができる。これらにより $|Y(t)| \neq 0$ が示された。同様に $|X(t)| \neq 0$ となる。 $\theta_w=0$ または $\theta_s=0$ の場合にも同じ結果が得られる。従って $X^*(t)$, $Y^*(t)$ の定義は意味を持つ。

4. 本手法のアルゴリズム

本手法の初期状態として、 N 個のデータを適当な平面に射影させたものを CRT 上に表示しておく。それ以後、次の各ステップを繰り返す (図 3)。

- 1° 現在 CRT 上に表示されているデータ点の座標を決定する 2 つの直交する p 次元単位ベクトルを X_1, Y_1 とする。すなわち、座標 $(X_1 \cdot x_a, Y_1 \cdot x_a)$ ($a=1, 2, \dots, N$) に点が表示されていると言える。
- 2° コンソールなどにより次に射影しようとする平面を、2 つの独立な p 次元ベクトルの成分により入力する。これらを直交する単位ベクトルに修正し、また必要に応じて補正する (後述)。この 2 つの直交する単位ベクトルを X_2, Y_2 とおく。
- 3° $t=0$ から 1 まで近似的に連続となる間隔で変化させ、それぞれの t に対して座標 $(X^*(t) \cdot x_a, Y^*(t) \cdot x_a)$ ($a=1, 2, \dots, N$) を CRT 上に表示する (古い点は消去してゆく)。

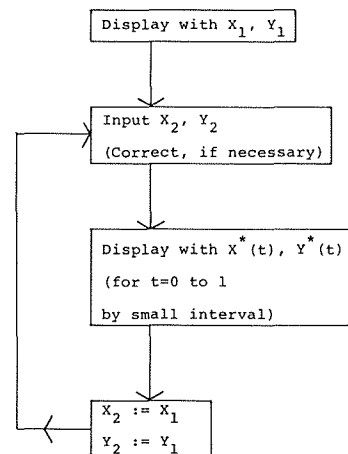


図 3 動的表示法のアルゴリズム

以上の説明では簡単のために省略したが、実際にはグラフィックディスプレイに合わせてデータのスケールなどを変更する必要がある。

上記のステップ 2° において行う射影する平面の指定は重要である。2 つのベクトルとして、次々

にランダムな値を入力してゆくのも多少効果があるが、原データの座標軸，主成分分析法から得られる主軸，P-P(Projection Pursuit) アルゴリズム⁸⁾の解，その他データに対して持っている情報の利用などを試みるほうがよい場合が多い。

5. 入力ベクトルの補正

前章のステップ 2°において触れた入力ベクトル（以下， X ， Y とおく）に対する補正などについて述べる。

(7)，(8)式によって定義された $X^*(t)$ ， $Y^*(t)$ を利用するためには直交した p 次元単位ベクトル X_2 ， Y_2 が必要である。そこで X ， Y に対して Schmidt の直交化を行ったものを X_2 ， Y_2 とする。また $X^*(t)$ ， $Y^*(t)$ の定義中でも述べたように $\pi/2 < \theta_w \leq \pi$ や $\pi/2 < \theta_s \leq \pi$ の場合には X_2 ， Y_2 の交換などを行う。以上により本手法はデータに対して適用可能となる。

しかし、応用によっては CRT 上の点の動きが少なくなるような補正を加えるほうがよい場合もある。入力ベクトル(直交性，正規性は仮定しておく) X ， Y に対して，同じ平面を張る 2 つの直交する p 次元単位ベクトルの集合において，

$$F(X_2, Y_2) = |X_1 - X_2|^2 + |Y_1 - Y_2|^2$$

が最小となる X_2 ， Y_2 を選ぶことを考える。これは

$$F(X_2, Y_2) = 4 - 2(X_1 \cdot X_2 + Y_1 \cdot Y_2)$$

より

$$G(X_2, Y_2) = X_1 \cdot X_2 + Y_1 \cdot Y_2$$

が最大となる X_2 ， Y_2 を選ぶことと同値である。

定理 2 $G(X_2, Y_2)$ が最大値をとるならば

$$\begin{cases} X_2 = \frac{\alpha + \delta}{\sqrt{(\alpha + \delta)^2 + (\beta - \gamma)^2}} X + \frac{\beta - \gamma}{\sqrt{(\alpha + \delta)^2 + (\beta - \gamma)^2}} Y \\ Y_2 = \frac{\gamma - \beta}{\sqrt{(\alpha + \delta)^2 + (\beta - \gamma)^2}} X + \frac{\alpha + \delta}{\sqrt{(\alpha + \delta)^2 + (\beta - \gamma)^2}} Y \end{cases} \quad (10)$$

(ただし， $(\alpha + \delta)^2 + (\beta - \gamma)^2 \neq 0$ のときに限る)

または，

$$\begin{cases} X_2 = \frac{\alpha - \delta}{\sqrt{(\alpha - \delta)^2 + (\beta + \gamma)^2}} X + \frac{\beta + \gamma}{\sqrt{(\alpha - \delta)^2 + (\beta + \gamma)^2}} Y \\ Y_2 = \frac{\beta + \gamma}{\sqrt{(\alpha - \delta)^2 + (\beta + \gamma)^2}} X + \frac{\delta - \alpha}{\sqrt{(\alpha - \delta)^2 + (\beta + \gamma)^2}} Y \end{cases} \quad (11)$$

(ただし， $(\alpha - \delta)^2 + (\beta + \gamma)^2 \neq 0$ のときに限る)

となる。ここで，

$$\alpha = X_1 \cdot X, \quad \beta = X_1 \cdot Y,$$

$$\gamma = Y_1 \cdot X, \quad \delta = Y_1 \cdot Y,$$

とする。

(証明) X と Y ， X_2 と Y_2 はそれぞれ直交する p 次元単位ベクトルなので

$$\begin{cases} X_2 = \cos t X + \sin t Y \\ Y_2 = -\sin t X + \cos t Y \end{cases} \quad (-\pi \leq t \leq \pi) \quad (12)$$

または

$$\begin{cases} X_2 = \cos t X + \sin t Y \\ Y_2 = \sin t X - \cos t Y \end{cases} \quad (-\pi \leq t < \pi) \quad (13)$$

とおける。また、

$$\begin{cases} X_1 = \alpha X + \beta Y + V \\ Y_1 = \gamma X + \delta Y + W \end{cases}$$

とする。ただし、

$$X \cdot V = Y \cdot V = X \cdot W = Y \cdot W = 0.$$

まず(12)式の場合について考察する。

$$\begin{aligned} G(X_2, Y_2) &= X_1 \cdot X_2 + Y_1 \cdot Y_2 \\ &= \alpha \cos t + \beta \sin t - \gamma \sin t + \delta \cos t \\ &= (\alpha + \delta) \cos t + (\beta - \gamma) \sin t \end{aligned}$$

ここで、 $\alpha + \delta \neq 0$ かつ $\beta - \gamma \neq 0$ のとき、

$$\begin{aligned} G(X_2, X_2) &= \sqrt{(\alpha + \delta)^2 + (\beta - \gamma)^2} \left\{ \frac{\alpha + \delta}{\sqrt{(\alpha + \delta)^2 + (\beta - \gamma)^2}} \cos t + \frac{\beta - \gamma}{\sqrt{(\alpha + \delta)^2 + (\beta - \gamma)^2}} \sin t \right\} \\ &= \sqrt{(\alpha + \delta)^2 + (\beta - \gamma)^2} \sin(t + \varphi) \end{aligned}$$

となる。ただし、

$$\sin \varphi = \frac{\alpha + \delta}{\sqrt{(\alpha + \delta)^2 + (\beta - \gamma)^2}}, \quad \cos \varphi = \frac{\beta - \gamma}{\sqrt{(\alpha + \delta)^2 + (\beta - \gamma)^2}}$$

$$-\pi \leq \varphi < \pi.$$

従って $G(X_2, Y_2)$ は

$$t + \varphi = \frac{\pi}{2}, -\frac{3\pi}{2}$$

のとき (12)式のもとでは) 最大となる。すなわち、

$$\begin{aligned} \sin t &= \cos \varphi \\ &= \frac{\beta - \gamma}{\sqrt{(\alpha + \delta)^2 + (\beta - \gamma)^2}} \end{aligned}$$

$$\therefore \cos t = \frac{\alpha + \delta}{\sqrt{(\alpha + \delta)^2 + (\beta - \gamma)^2}}$$

となり(10)式が成立する。これは一般に $(\alpha + \delta)^2 + (\beta - \gamma)^2 = 0$ でない限り成立する。同様にして(13)式の場合には(11)式が成立することが示される。 **Q. E. D.**

注意：本定理では $(\alpha + \delta)^2 + (\beta - \gamma)^2 = 0$ かつ $(\alpha - \delta)^2 + (\beta + \gamma)^2 = 0$ の場合、すなわち $\alpha = \beta = \gamma = \delta = 0$ の場合には触れていないが、このときには $G(X_2, Y_2) \equiv 0$ となる。この定理 2 により $G(X_2, Y_2)$ を最大とする、 X_2, Y_2 を選ぶことができる。

6. 適用例

R. A. Fisher (1936)⁷⁾による有名な Iris データに対して本手法を適用した例を示す。ただし、『点の動き』を紙上に直接、画くことは不可能なので点の軌跡で代用する。また P-P アルゴリズムの解などは Friedman & Tukey (1974)⁸⁾の値を利用した。

Iris データはがくの長さ、がくの幅、花卉の長さ、花卉の幅を変量にしたもので、setosa, versicolour, virginica が各 50 個づつ、計 150 個からなるものである。ここでは平均が 0 となるように変換した

値を用いる。

図4では150個のデータに対して X_1 は第1成分の値, Y_1 は第2成分の値, X と Y は全データに対する P-P アルゴリズムによる解としたものである。図5は *setosa* を除く100個に対して X_1 , Y_1 は主成分分析から得られる主軸, X と Y は P-P アルゴリズムによる解としたものである。これは X_2 , Y_2 に対する補正を加えていないが, $G(X_2, Y_2)$ が最大となるような補正(前述)を加

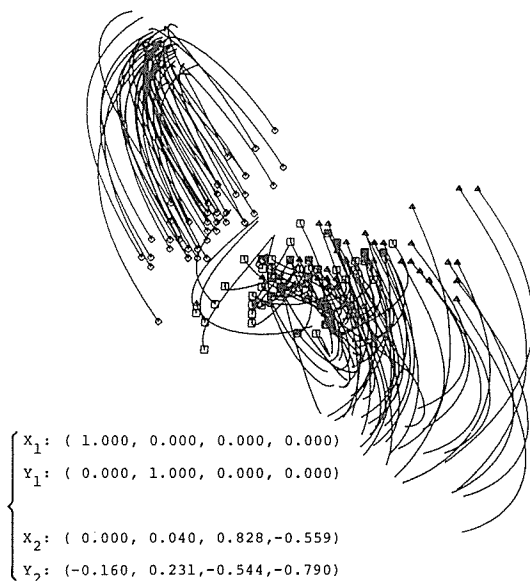


図4 Iris データによる軌跡

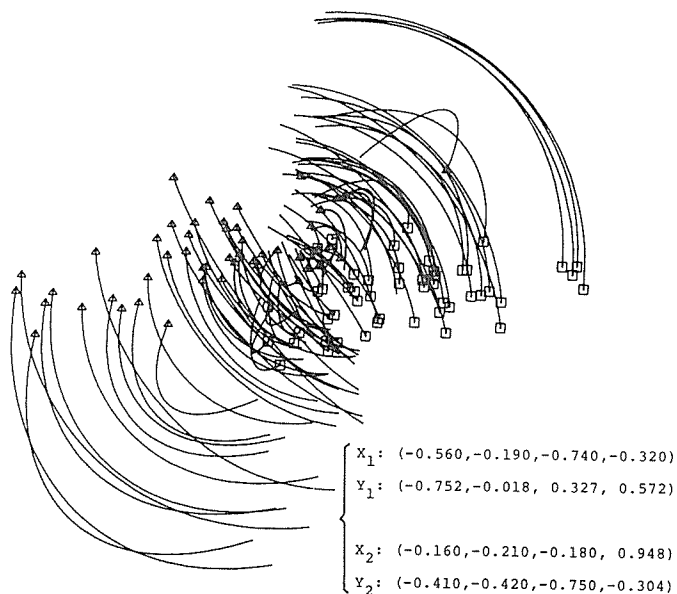


図5 *setosa* を除く Iris データによる軌跡

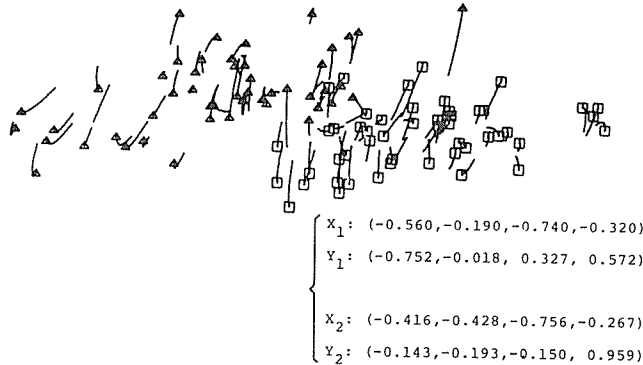


図 6 補正を加えた軌跡

えると図 6 のような結果となる。

特に図 6 からは、主成分分析による結果と P-P アルゴリズムの解との関係を調べる一つの資料が得られた。

7. 本手法の特徴および結論

本手法には以下の特徴がある。

- 1) CRT 上には点しか表示していないので、データの個数がある程度多くてもよい。
- 2) 表示されている画面は常にデータの線形な射影なので表示に客観性があり、結果の解釈が行い易い。
- 3) アルゴリズムが単純なので比較的安価なマイコン等でも実現できる。
- 4) 射影する方向を指定する自由性があるが、反面その指定が煩わしい場合もある。
- 5) 他のデータ解析の手法（主成分分析，因子分析，クラスター分析，判別分析等）と結合させることによりシステム化し易い。

特に 5) のシステム化については重要で、本手法そのものはデータ解析における補助手段の一つとみるべきである。そこで他の手法と結合させたり、射影方向の入力方法や点のカラー化について実際の応用に合わせて工夫すべきであろう。

グラフィックディスプレイが安価で高性能になりつつあるので、本手法のように多次元データを動的に表示することは容易になってきた。そこで動的表示を用いたグラフ表現の一層の研究は意味があると思われる。また、これ以外にもパーソナルコンピュータを、データ解析に対して効果的に用いる方法の開発は必要である。

参 考 文 献

- 1) Sammon J. S.: IEEE Trans. on Computers, C-18 (1969), p. 401-409.
- 2) Chernoff, H.: J. Amer. Statist. Assoc., 68 (1973), p. 361-368.
- 3) Wakimoto, K. & Taguri, M.: Ann. Inst. Statist. Math., 30 Part A (1978) p. 77-84.
- 4) Andrews, D. F.: Biometrics, 28 (1972), p. 125-136.
- 5) 新保 勝・宮腰政明: 工学部研究報告, 96 (1979) p. 65-71.
- 6) Andrews, D. F.: International Roundtable Congress (In Tokyo) (1981)
- 7) Fisher, R. A.: Ann. Eigen., 7 (1936) p. 179-188.
- 8) Friedman, J. H. & Tukey, J. W.: IEEE Trans. on Computers, C-23 (1974) p. 881-890.

- 9) 水田正弘・斉藤正明・佐藤義治・河口至商：多変量解析研究会研究報告，16（1980）p. 1-6.
- 10) 水田正弘：マイクロコンピュータ札幌シンポジウム論文集，3（1980）p. 58-59.