Support Vector Machine を利用した大分類関数の構成

鈴木　昇一

A Construction of a Rough Classifier Having a Support Vector Machine as Its Structure

Shoichi Suzuki

あらまし

SS理論と名付けられたパターン認識の数学的理論に登場する RECOGNITRON は、処理の対象とする問題の入力パターン φ に対応し、"axiom 1を満たすパターンモデル" $T\varphi$ を求め、$T\varphi$ を恰も、φ かのように扱う。このとき、写像 T はパターンモデル構成作用素と呼ばれる。axiom 2, 3を各々満たす類似度関数 SM, 大分類関数 BSC を構成すれば、RECOGNITRON は φ に関する連想形認識方程式を解くことによって、φ から連想されるパターンと、φ の帰属するカテゴリを求めることができる。

本論文では、あるカテゴリに帰属するか否かに分類される訓練データに関し、2分割された訓練データ間のマージンが最大になるような超平面を求める2カテゴリ学習分類器、サポートベクタマシン（SVM）の理論を適用し、axiom 3を満たす大分類関数 BSC を設計する手法が提案される。

計算論的学習理論の1つとしての "適応的ブースティングアルゴリズム Ada Boost" を適用して、BSC を設計できることは既に示されている。完全に線形分離でなくても分類誤差を考慮に入れて分離境界を与える超平面を決定するSVM理論の適用により、BSC の設計が訓練パターン集合について適切に設計できる1つの手法が得らている。

キーワード

パターン認識の数学的理論（SS理論）　モデル構成作用素　類似度関数
大分類関数　2次計画問題　サポートベクタ

Abstract

RECOGNITRON appearing in a mathematical theory of recognizing patterns named SS-theory seeks from an input original pattern φ in question to be recognized a corresponding pattern-model $T\varphi$ which must satisfy axiom 1 suggested by S.Suzuki, and treats $T\varphi$ as though $T\varphi$ would be φ. The mapping T is called a model-construction operator. Provided that a similarity-measure function SM and a rough classifier BSC are constructed so as to respectively satisfy axiom 2 and axiom 3, RECOGNITRON can determine a pattern recalled from pattern φ and a category to which φ belongs by solving an associative equation of recognition about φ.

In this paper, a BSC is designed according to a theory of support-vector machine(SVM). A learning machine SVM which can divide into two subsets of patterns seeks for two hyperplanes whose margin is maximized for a training set of patterns.

It was evident that BSC could be designed by applying a boosting algorithm Ada Boost in a computational learning theory. SVM has an ablity of determining two hyperplanes which give two boundaries considering an error of classification whatever the traing set may not be linearly separable. Therefore a method of designing BSC is obtained with the object of adjusting RECOGNITRON to the training set.
Key words : a mathematical theory of recognizing patterns (SS theory)　model-construction operator　similarity-measure function　rough classifier　rough classifier　quadratic programming problem　support vector

1. まえがき

認識システム RECOGNITRON [B3], [B4] は、処理の対象とする問題のパターンφのカテゴリ帰属知識に関する連想形認識方程式を解くことにより、φから連想されるパターンと、φの帰属するカテゴリを求めるが、この連想形認識方程式は、3axiom 1, 2, 3を各々満たすモデル構成作用素 T, 類似度関数 SM, 大分類関数 BSC を構成すれば、決まる。

計算論的学習理論の1つとしての“適応的ブースティングアルゴリズム Ada Boost”を適用して、BSC を設計できることは既に示されている [B24]。

本論文は、高次元の特徴量の組を入力として扱え、然も過学習を起こさずに最適解を求めることのできるSVM理論を適用し、axiom 3を満たす大分類関数 BSC を設計する手法を研究したものである。

例えば、動画像を用いて遠隔地の現在状況を把握する施設監視システムでは、コンピュータが画像、映像や音声などの非言語情報を処理することになる。このために、画像(内容を)理解(する)システムが動画像中の人物が行う動作などを抽出し、その結果をコンピュータ(知能情報メディア)は簡潔な自然言語で説明する必要がある。このように、非言語情報と言語情報を結び付けるマルチメディア技術が使われている。同様に、人間同士間の自然な会話をコンピュータに認識理解させ、文字化させる“大語彙連続音声認識理解システム”の開発は、マルチメディア社会にとって必要なものである。

マルチメディア社会で取り扱われる情報は、文字列(で表される)言語と、パターン(で表される)時系列である。

マルチメディア時代の入り口を通過した現在、

(1) メディアで表された情報を検索し、認識･理解する技術の確保問題

(2) 例えば、テキスト(記号列；文字列)から音声･画像への変換といったメディア変換技術の確立問題

(3) 人間機能を代行しながら、知的に振る舞う知的エージェントの構成問題
などを解決しながら、多種多様なメディアを益々、高効率に処理しなければならない。このために基本的に要求されるのが、自然言語(テキスト)の処理技術(言語による表象(命題表象［A17］)の処理技術)、パターン列の処理技術(視覚・聴覚などによる表象(アナロジー表象［A17］)の処理技術)である。

パターンとは、静止画像、動画像、平面画像、立体画像、言語音声、会話音声等などの総称である。

パターン情報学では、分類の対象となるものをパターンというが、S.Suzuki以外のパターン想起・認識の理論は，パターンが既に圧縮されたものをパターンと称して論が展開されることが多い.既に圧縮されているものは事実上そのパターンの特徴量の組であるにもかかわらず。分類の前処理としてなされるデータ圧縮はこの意味で、ある程度似た者同士を1つにまとめるクラスタリングの働きをしている。

処理の対象とする問題のパターン $\varphi \in \Phi$ を圧縮したものがS.Suzuki理論でのパターンモデル $T\varphi \in \Phi$ である。$T\varphi$ を見たり聞いたりしたならば、φ であるかのように見えたりするためには，処理の対象とする問題のパターンの集合 Φ と、写像

$$T : \Phi \rightarrow \Phi \tag{1.1}$$

との対【Φ, T】は、少なくともaxiom 1を満たさなければならないというのがS.Suzuki理論［B1］～［B4］の主張である。このような写像 T はモデル構成作用素と呼ばれる。

S.Suzukiは、表象化・知覚・連想・記憶・検索・認識・学習・理解に関するパターン情報処理の知能的問題解決理論を

“axiom 1～axiom 4の4公理からなるSS公理系から導かれる
パターン認識の数学的理論(SS理論)［B1］～［B6］” (1.2)

を拠り所として確立しようとしている。ここに、例えば、外界の状況を知識(長期記憶内容)を用いての、何らかの推論(連想)の働きで再構成しながら、知識に基づいて外界(の各対象と、それらの間の相互関係)を意味付けすることが、(外界)**理解**である。

S.Suzuki理論を適用し、外界を理解する能力を備えたシステムを現実場面で活用するには、axiom 1, 2, 3を各々満たす式(1.1)のモデル構成作用素 T，類似度関数

$$SM : \Phi \times \Omega \rightarrow \{s \mid 0 \leqq s \leqq 1\} \tag{1.3}$$

，並びに、大分類関数

$$BSC : \Phi \times J \rightarrow \{0, 1\} \tag{1.4}$$

の3者を具体的に設計しなければならない。これまでT, SM, BSCについては、文献Bに見られるごとく、それらの具体的な設計論はある程度研究されてきた。ここに、Ωは式(2.19)での代表パターン集合であり、J はカテゴリ番号の集合である。

本研究論文の目的は、新たに、SVM(support vector machine)［A4］～［A7]，［A20]，［A21］によって、axiom 3を満たす式(1.4)の大分類関数 BSC を設計すること、つまり、

a method of constructing BSC from empirical data

を提供することである(新規性)。

SVMとは、特徴量の組が低次元超平面によって線形分離可能でないとき，高次元超平面によって線形分離可能にすることを目的として設計される学習機械(learning machine)のことである。SVMはC.Cortes, V.Vapnik［A20］によって提案された“重みベクトルが support vectors を陽に持

つ1次結合によって与えられる2カテゴリ分類のための1次識別関数" である。

SVMへの入力は通常、パターンから抽出された特徴量の組であるから、パターン $\varphi \in \Phi$ から抽出された第 $\ell \in L$ 番目の特徴量 $u(\varphi, \ell) \in R$(実数全体の集合)の組

$$\underline{u}(\varphi) = \{u(\varphi, \ell) \mid \ell \in L\} \in Rq \text{(q次元実数値の集合)} \quad (1.5)$$

を使って、式(1.4)の BSC を構成することになる。ここに、特徴抽出写像

$$u : \Phi \times L \rightarrow R \quad (1.6)$$

が導入されていることに注意しておく。

SVM は通常、線形分離ではない $\underline{u}(\varphi)$ の集合に対しても、余裕を持って、誤分類をできるだけ少なく2カテゴリの境界面を設定する機能を備えている故に、設計された BSC は現実の適用場面において有効である。それのみならず、評価の定まったSVM理論の応用として、BSC を設計した故に本研究内容の信頼性は保証されているといってもよかろう。

2. 処理の対象となる問題のパターン φ の集合 Φ とモデル構成作用素 T との対【Φ, T】と、類似度関数 SM

本章では，処理の対象となる問題のパターン φ の集合 Φ，モデル構成作用素 T，類似度関数 SM について説明される。対【Φ，T】の満たさなければならないaxiom 1と、SM の満たさなければならないaxiom 2も説明され、Φ の表示、T， SM の構成例が示される。

2.1 処理の対象となる問題のパターン φ の集合 Φ，モデル構成作用素 T と、axiom 1を満たす対【Φ，T】

認識システムRECOGNITRONがモデル$T\varphi \in \Phi$を見たり聞いたりしたならば、原パターン$\varphi \in \Phi$と同じに見えたり聞こえたりすることだと、解釈可能な対【Φ，T】について説明しよう。

Φ は処理の対象とする問題のパターン φ の集合であり、$T\varphi \in \Phi$ は $\varphi \in \Phi$ に対応するパターンモデルであって、$T\varphi \in \Phi$ を見たり聞いたりしたならばあたかも原パターン $\varphi \in \Phi$ かのように見えたり聞こえたりするようなものである。このとき、モデル構成作用素(model-construction operator)

$$T : \Phi \rightarrow \Phi \quad (2.1)$$

が導入される。

SS理論では，対【Φ，T】はaxiom 1を満たしていなければならない。

一般に，処理の対象とする問題のパターンφの集合Φを可分な(separable)一般抽象ヒルベルト空間$\mathfrak{H}$の或る部分集合とすると、パターンモデル$T\varphi$を出力する式(2.1)の写像 T に要求されるのは、次の4性質①～④であることが理論的に明らかにされている [B3]，[B4] ：

①(零元不動点性) $\varphi = 0 \in \Phi$については、$T\varphi = 0$.

②(正定数倍不変性)任意の正実定数aに対し,

$\forall \varphi \in \Phi, T(a \cdot \varphi) = T\varphi$.

③(ベキ等性) $\forall \varphi \in \Phi, T(T\varphi) = T\varphi$.

④(非零写像性) $\exists \varphi \in \Phi, T\varphi \neq 0$. □

尚、$\overline{\eta}$をηの複素共役として,

M：q次元ユークリッド空間R^qの可測部分集合 (2.2)

dm(x)：正値ルベーグ･スティルチェス式測度 (2.3)

$x = \langle x_1, x_2, \cdots, x_q \rangle \in M(\subseteq R^q)$：実数値多座標変数 (2.4)

を導入し、その内積(φ, η)、ノルム$\|\varphi\|$を

$$(\varphi, \eta) = \int_M dm(x)\,\varphi(x)\cdot\overline{\eta}(x) \tag{2.5}$$

$$\|\varphi\| \equiv \sqrt{(\varphi, \varphi)} \tag{2.6}$$

とする線形空間(ベクトル空間)としての可分なヒルベルト空間[19] $\mathfrak{H} = L_2(M；dm)$の特別な場合として、

$$M = R^2 \text{(2次元全平面)} \tag{2.7}$$

$$dm(x) = [x_1^2 + x_2^2]^{-1} dx_1 dx_2 \tag{2.8}$$

を選ぶことができる［B7]，[B9]。

処理の対象とするパターンφの集合Φはある可分なヒルベルト空間[19] $\mathfrak{H}$の、零元0を含むある部分集合であり、このΦ、並びに式(2.1)の写像Tの対【Φ, T】は2.の4性質①～④((ⅰ),(ⅱ),(ⅲ)の3後半，並びに(ⅳ))を含む形で，次のaxiom 1を満たさなければならない。このとき，写像Tはモデル構成作用素と呼ばれ、$T\varphi \in \Phi$は$\varphi \in \Phi$の代りとなり得るという意味で、パターン$\varphi \in \Phi$のモデル(model)と呼ばれる。

Axiom 1(パターン集合Φとモデル構成作用素Tとの対【Φ, T】の満たすべき公理)［B3]，[B4]

(ⅰ)(零元0のΦへの埋込性，零元0のT－不動点性)

$$0 \in \Phi \wedge T0 = 0.$$

(ⅱ)(Φの錐性，Tの正定数倍吸収性)

$$\forall \varphi \in \Phi, a\cdot\varphi \in \Phi \wedge T(a\cdot\varphi) = T\varphi$$

for any positive real number a.

(ⅲ)(Φへの埋込性，Tのベキ等性)

$$\forall \varphi \in \Phi, T\varphi \in \Phi \wedge T(T\varphi) = T\varphi.$$

(ⅳ)(Tの非零写像性)

$$\exists \varphi \in \Phi, T\varphi \neq 0.$$ □

上述のaxiom 1からわかるように、処理の対象とする問題のパターン集合Φは、埋込性

$$T\cdot\Phi \equiv \{T\varphi \mid \varphi \in \Phi\} \subset \Phi \tag{2.9}$$

を満たし，原点(＝0)を始点とし、Φの任意の点を通る半直線を含むような集合、つまり、錐(cone)であらねばならない。

パターンと判明しているφの集合(基本領域；basic domain) $\Phi_B(\ni 0)$と、すべての正実定数の集合 R^{++}とを用意する。

次の定理2.1は、axiom 1を満たす対［Φ, T］を決定している。

［定理2.1］(パターン集合Φとモデル構成作用素Tとの対［Φ, T］の基本構成定理)

式(2.1)の写像Tがaxiom 1の(ⅰ)，(ⅱ)，(ⅲ)の3後半、並びに、(ⅳ)を満たすとしよう。このとき、次の(イ)、(ロ)が成り立つ：

(イ)処理の対象とする問題のパターンの集合Φを、

$$\begin{aligned}\Phi &= R^{++}\cdot(\Phi_B \cup T\cdot\Phi_B)\\ &\equiv \{r^{++}\varphi_B \mid r^{++} \in R^{++},\ \varphi_B \in \Phi_B\}\\ &\quad \cup \{r^{++}T\varphi_B \mid r^{++} \in R^{++},\ \varphi_B \in \Phi_B\}\end{aligned} \tag{2.10$_1$}$$

の如く設定すれば、

$$\Phi \supset \{0\} \wedge R^{++} \cdot \Phi = \Phi \wedge [T \cdot \Phi = T \cdot \Phi_B \subset \Phi] \tag{2.11}$$

が成立し、axiom 1の(i)、(ii)、(iii)の3前半をΦは満たし、結局、対【Φ, T】はaxiom 1を満たす。

(ロ)逆に、$\Phi_B(\ni 0)$を部分集合に持つΦがaxiom 1の(i)、(ii)、(iii)の3前半を満たすとすれば、

$$\Phi \supseteq \Phi_B \cup R^{++} \cdot \Phi \cup T \cdot \Phi \tag{2.10_2}$$

と表されるが、ここで、特に包含式(2.10_2)において等号が成立するような最小のΦを採用すれば、axiom 1を満たす対【Φ, T】のΦは式(2.10_1)のように表され、式(2.11)も成立する。

(証明)(イ)は文献［B4]，付録1の定理A1.1である.(ロ)は文献［B3]，pp.64-66(2.4節)で証明されている. □

2.2 モデル構成作用素 T の構成例

モデル構成作用素 T を1つ、構成しておこう。

【axiom 1の(i)、(ii)、(iii)の3後半、並びに、(iv)を満たすモデル構成作用素 T の構成例】

本例では，可分なヒルベルト空間 $\mathfrak{H} = L2(M\,;\,dm)$を採用し，処理の対象とする問題のパターン$\varphi$の集合Φを$\Phi \subset L_2(M\,;\,dm)$とする.

可測部分集合$(x \in)M(\subseteq Rq)$を考え，

$$\forall x \in M,\ (S\varphi)(x) = \begin{cases} 0 \cdots \sup_{x\in M} |\varphi(x)| = 0\text{のとき} \\ \varphi(x)/\sup_{x\in M} |\varphi(x)| \cdots \sup_{x\in M} |\varphi(x)| > 0\text{のとき} \end{cases} \tag{2.12}$$

と定義される写像

$$S : \Phi \to \Phi \tag{2.13}$$

を導入し，不等式

$$\forall x \in M,\ 0 \leqq h(x) < 1 \tag{2.14}$$

を満たす閾値関数 h(x)を導入すると，

$$(T\varphi)(x) = \begin{cases} 0 \cdots (S\varphi)(x) \leqq h(x)\text{のとき} \\ 1 \cdots (S\varphi)(x) > h(x)\text{のとき} \end{cases} \tag{2.15}$$

と定義される式(2.1)の写像Tは1.の4性質①～④を満たす.

第$j \in J$番目のカテゴリ$\mathfrak{C}_j$の生起確率$p(\mathfrak{C}_j)$と、$\mathfrak{C}_j$の代表パターンω_jとを設けると、その平均化パターンξは、

$$\xi = \sum_{j\in J} p(\mathfrak{C}_j) \cdot \omega_j \tag{2.16}$$

と定義される。

1より小さい十分小さい正値関数$\varepsilon(x)$を，不等式

$$\forall x \in M,\ 0 < (S\xi)(x) < 1 \Rightarrow (S\xi)(x) < 1 - \varepsilon(x) \tag{2.17}$$

を満たすように導入し、更に、$\xi(x)$を式(2.16)の平均化パターンとすれば、閾値関数 h(x)として、

$$h(x) = \begin{cases} 1 - \varepsilon(x) \cdots (S\xi)(x) = 1\text{の場合} \\ (S\xi)(x) \cdots 0 \leqq (S\xi)(x) < 1\text{の場合} \\ 0 \cdots (S\xi)(x) \leqq 0\text{の場合} \end{cases} \tag{2.18}$$

を採用したこのパターンモデル$T\varphi$は、文献［B17］で顔画像φの2値化画像を得るために使われている。

訓練パターン系列を設け、この系列からの学習で閾値関数$h(x)$を適切に決定すれば、$T\varphi$はφの骨格を表す.このようにして、原パターンφの骨格を表すパターンモデル$T\varphi$(2値化パターンモデル)が得られたことになる。

2.3 代表パターン集合 Ω

第$j \in J$番目のカテゴリ$\mathfrak{C}_j$の持つ諸性質を典型的に代表しているパターンを代表パターンと呼び、

$$\omega_j \in \Omega \equiv \{\omega_j \mid j \in J\} \tag{2.19}$$

と表そう。

Ωを視察で決定できる場合もあるが、訓練パターン系列からΩを適応的に決定する方法については、文献［B3］の付録Iで説明されている。

尚、式(2.16)で登場している非負実数$p(\mathfrak{C}_j)$は、2条件

$$[\forall j \in J, 0 < p(\mathfrak{C}_j) < 1] \wedge [\sum_{j \in J} p(\mathfrak{C}_j) = 1] \tag{2.20}$$

を満たしていなければならない。

2.4 axiom 2を満たす類似度関数 SM

$$SM(\varphi, \omega_j) \in \{s \mid 0 \leqq s \leqq 1\} \tag{2.21}$$

は、パターン $\varphi \in \Phi$ がω_jと似ている程度を表す類似度であって、類似度関数(similarity-measure function)

$$SM: \Phi \times \Omega \to \{s \mid 0 \leqq s \leqq 1\} \tag{2.22}$$

が導入される。

SS理論では、類似度関数 SM は次のaxiom 2を満たしていなければならない。

Axiom 2(類似度関数 SM の満たすべき公理)［B3],［B4］

(i)(正規直交性) $\forall i, \forall j \in J, SM(\omega_i, \omega_j) = \delta_{ij}$.

(ii)(規格化性) $\forall \varphi \in \Phi, \sum_{j \in J} SM(\varphi, \omega_j) = 1$.

(iii)(写像Tの下での不変性)

$\forall \varphi \in \Phi, \forall j \in J, SM(T\varphi, \omega_j) = SM(\varphi, \omega_j)$. □

上述のaxiomの(i)では、クロネッカーのδ記号

$$\delta_{ij} = 1 \text{ if } i = j, = 0 \text{ if } i \neq j \tag{2.23}$$

が導入されている。

上述のaxiomの(i)〜(iii)について簡単に説明しておこう。

$SM(\varphi, \omega_j) = 1, 0$に従って、パターン$\varphi \in \Phi$は各々、

ω_jと確定的な類似関係、相違関係にあり、また,

$0 < SM(\varphi, \omega_j) < 1$の場合は、曖昧な類似·相違関係にある (2.24)

と，SMを解釈しよう。(i)は、相異なるカテゴリの代表パターン同士は確定的な相違関係にあり、同一カテゴリの代表パターン同士は確定的な類似関係にあることを要請している。(ii)は、任意のパターンφについて，すべてのカテゴリについての類似度の総和は1であることを要請している。(iii)は、パターンモデル$T\varphi$は原パターンφと任意のカテゴリについて同一類似度を持つことを要請している。ということは、パターンモデル$T\varphi$を見たり，聞いたりするならば，原パタ

ーンφと同じように見えたり、聞こえたりすること(**同一知覚原理**)を要請していることになる。

これまで，上述のaxiomを満たす類似度関数SMは多数構成されており［B3］，［B4］，［B13］，［B14］，［B17］，［B19］～［B22］，その有効性についても計算機シミュレーション済［B13］，［B14］，［B17］，［B20］である．

2.5 類似度関数 SM の構成例

axiom 2を満たす式(2.22)の類似度関数 SM を1つ、構成しておこう。

2条件

$$\forall j \in J,\ T\omega_j \in T\Psi_j \equiv \{T\psi \mid \psi \in \Psi_j\} \tag{2.25}$$

$$\forall j \in J,\ \forall i \in J-\{j\},$$
$$T\Psi_i \cap T\Psi_j = \phi \text{ (the empty set)} \tag{2.26}$$

の下で、関数 $g_j(\varphi)$を、

$$g_j(\varphi) = \min_{\psi \in \Psi_j} \left[1 - |(T\varphi \| T\varphi \|^{-1}, T\psi \| T\psi \|^{-1})|^2\right] \tag{2.27}$$

と定義する。ここに、

$$(T\varphi \| T\varphi \|^{-1}, T\psi \| T\psi \|^{-1}) = 0 \text{ if } \| T\varphi \| \cdot \| T\psi \| = 0 \tag{2.28}$$

と、約束している。

$$\forall j \in J,\ \varphi \in \Psi_j \Rightarrow g_j(\varphi) = 0 \tag{2.29}$$

$$\forall j \in J,\ \forall i \in J-\{j\},\ \varphi \in \Psi_j \Rightarrow 1 \geqq g_i(\varphi) > 0 \tag{2.30}$$

が成立している。その後、関数 $f_j(\varphi)$を、

$$f_j(\varphi) = \min_{i \in J-\{j\}} g_i(\varphi) \tag{2.31}$$

と定義すると、

$$\forall j \in J,\ \varphi \in \Psi_j \Rightarrow f_j(\varphi) > 0 \tag{2.32}$$

$$\forall j \in J,\ \forall i \in J-\{j\},\ \varphi \in \Psi_j \Rightarrow f_i(\varphi) = 0 \tag{2.33}$$

が成立している。よって、

$$SM(\varphi, \omega_j) = \begin{cases} f_j(\varphi) / \sum_{i \in J} f_i(\varphi) & \cdots \sum_{i \in J} f_i(\varphi) > 0 \text{ のとき} \\ p(\mathfrak{C}_j) & \cdots \sum_{i \in J} f_i(\varphi) = 0 \text{ のとき} \end{cases} \tag{2.34}$$

と定義される式(2.22)のSMは、

$$\forall j \in J,\ \varphi \in \Psi_j \Rightarrow SM(\varphi, \omega_j) = 1 \tag{2.35}$$

$$\forall j \in J,\ \forall i \in J-\{j\},\ \varphi \in \Psi_j \Rightarrow SM(\varphi, \omega_i) = 0 \tag{2.36}$$

を満たし、axiom 2の(i)を満たすことがわかる。axiom 2の(ii)，(iii)をも満たし、結局、axiom 2を満たす。

尚、不等式

$$\forall j \in J,\ 0 \leqq s_0(j) < s_1(j) \leqq 1 \tag{2.37}$$

を満たす2つの閾値$s_0(j)$，$s_1(j)$を見つけることができる。このとき，axiom 2を満たす類似度関数$SM'(\varphi, \omega_j)$を

$$s(\varphi,\ \omega_j)=$$
$$\begin{cases} 1 & \cdots s_1(j) \leqq SM'(\varphi,\ \omega_j)\text{の場合} \\ [SM'(\varphi,\ \omega_j)-s_0(j)]/[s_1(j)-s_0(j)] & \\ \quad \cdots s_0(j) < SM'(\varphi,\ \omega_j) < s_1(j)\text{の場合} & \\ 0 & \cdots SM'(\varphi,\ \omega_j) \leqq s_0(j)\text{の場合} \end{cases} \tag{2.38}$$

へと，区分的線形変換を使い変換すると，

$$SM(\varphi,\ \omega_j)=$$
$$\begin{cases} s(\varphi,\ \omega_j)/\sum_{i\in J} s(\varphi,\ \omega_i) & \\ \qquad \cdots \sum_{i\in J} s(\varphi,\ \omega_i) > 0\text{の場合} & \\ p(\mathfrak{C}_j)\cdots \sum_{j\in J} s(\varphi,\ \omega_j)=0\text{の場合} \end{cases} \tag{2.39}$$

と定義される非負実数値関数関数SMはSM´の性質を受け継いでおり，axiom 2を満たすことがわかる。

2.6　類似度関数 SM から眺めた処理すべきパターン集合 Φ と、構成しなければならない認識システムが備えていなければならない認識性能

パターンと称されてよい学習すべき各基本パターンの、ごく近くにあるものの集まりが、処理すべき問題のパターン φ の集合 Φ であると考えられ、不等式

$$0 \leqq \delta_j < 2^{-1} \tag{2.40}$$

を満たす或る非負実数 δ_j を考えると、このパターン集合 Φ は互いに素な集合

$$\Phi_j=\{\varphi\in\Phi \mid SM(\varphi,\ \omega_j) \geqq 1-\delta_j\} \tag{2.41}$$

の和と、どのカテゴリにも帰属しないパターンや2つ以上のカテゴリに帰属しているパターンとの集まりである Φ_0 との和として、

$$\Phi=\bigcup_{j\in J} \Phi_j \cup \Phi_0 \tag{2.42}$$

と表されると考えられる。

式(2.41)で表されるパターン集合 Φ_j の各元 φ については第 $j\in J$ 番目のカテゴリ $\mathfrak{C}_j$ に帰属する出力をもたらす認識システムを構成しなければならないことは、次の定理2.2からわかる。

［定理2.2］（**SM－δ_j 定理**）

（i）（一意的帰属に関するSM－δ_j 定理）

不等式

$$SM(\varphi,\ \omega_j) \geqq 1-\delta_j \tag{2.43}$$

を満たすカテゴリ番号 $j\in J$ は存在するとすれば、唯1つしかない。

（ii）（一意的帰属に関するSM－maxの 2^{-1} 分離定理）

不等式(2.43)が成立していれば、

$$\max_{i\in J-\{j\}} SM(\varphi,\ \omega_i) \leqq \delta_j < 2^{-1}$$
$$< 1-\delta_j \leqq SM(\varphi,\ \omega_j). \tag{2.44}$$

（iii）（一意的帰属に関するSM－maxの差 $1-2\delta_j$ 定理）

不等式(2.43)が成立していれば、

$$0 < 1-2\delta_j \leqq SM(\varphi,\ \omega_j) - \max_{i\in J-\{j\}} SM(\varphi,\ \omega_i). \tag{2.45}$$

（証明）　（i）の証明：不等式(2.43)を満たすカテゴリ番号 $i\in J_1\subseteq J$ が $|J_1| \geqq 2$ 個、存在するとしよう。

$$1 = \sum_{j \in J - J_1} SM(\varphi, \omega_j) + \sum_{i \in J_1} SM(\varphi, \omega_i)$$

$\because$ axiom 2の(ii)

$$\geqq \sum_{i \in J_1} (1 - \delta_j)$$

$$> | J_1 | \cdot 2^{-1} \qquad \because \text{式}(2.40)\text{から } 1 - \delta_j > 2^{-1}$$

$$\geqq 1 \tag{2.46}$$

を得、これは矛盾である。

(ii)の証明：

$$1 = \sum_{i \in J} SM(\varphi, \omega_i) \qquad \because \text{ axiom 2の(ii)}$$

$$= SM(\varphi, \omega_j) + \sum_{i \in J - \{j\}} SM(\varphi, \omega_i)$$

$$\geqq SM(\varphi, \omega_j) + \max_{i \in J - \{j\}} SM(\varphi, \omega_i)$$

$\therefore$

$$1 - SM(\varphi, \omega_j) \geqq \max_{i \in J - \{j\}} SM(\varphi, \omega_i) \tag{2.47}$$

であるが、不等式(2.43)から

$$\delta_j \geqq 1 - SM(\varphi, \omega_j)$$

が成立しているから、

$$\max_{i \in J - \{j\}} SM(\varphi, \omega_i) \leqq \delta_j \tag{2.48}$$

が成立する。残りは2式(2.40), (2.43)から明らかである。

(iii)の証明： 不等式(2.48)から得られる不等式

$$- \max_{i \in J - \{j\}} SM(\varphi, \omega_i) \geqq - \delta_j$$

と、不等式(2.43)とを加えれば、

$$SM(\varphi, \omega_j) - \max_{i \in J - \{j\}} SM(\varphi, \omega_i) \geqq 1 - 2\delta_j$$

$$> 0 \tag{2.49}$$

を得、不等式(2.45)が得られた。 □

2.7 パターンモデルTφを不変に保つパターン変換Uからもたらされる類似度関数 SM の不変性

モデル構成作用素 T があるパターン変換 U に対し不変ならば、類似度関数 SM もパターン変換 U に対し不変であることを説明しよう。

パターン変換

$$U : \Phi \to \Phi \tag{2.50}$$

に関し、等式

$$T(U\varphi) = T\varphi \tag{2.51}$$

が成立するならば、パターンモデルTφを生成する式(2.1)の写像Tは、パターンφの変形

$$\varphi \to U\varphi \tag{2.52}$$

を吸収する能力を備えている。何故ならば、φとその変形Uφは共に、共通なパターン標準形Tφを持つことになるからである。この種のパターン変換として規則的変形としてのユニタリ座標変換、不規則的な変形をを許容する離散量子化変換があることは既に示されている [B1], [B3] ~ [B5]。

類似度関数SMの2種類の不変性について説明しよう。

不変性

$$\forall \varphi \in \Phi, T(U\varphi) = T\varphi \tag{2.53}$$

を満たす任意のパターン変換Uは大抵の場合、多数存在する［B1］,［B5］,［B9］,［B10］,［B18］.例えば，axiom 1,（ii）の後半での、任意の正実数 a がそうである。このとき、

（イ）(SMの正定数倍不変性)任意の正定数 a について、

$$\forall \varphi \in \Phi,\ \forall j \in J,$$

$$SM(a \cdot \varphi,\ \omega_j) = SM(\varphi,\ \omega_j) \tag{2.54}$$

が成立する。何故ならば、

$$\begin{aligned} & SM(a \cdot \varphi,\ \omega_j) \\ & = SM(T(a \cdot \varphi),\ \omega_j) \quad \because \text{ axiom 2の(iii)} \\ & = SM(T\varphi,\ \omega_j) \quad \because \text{ axiom 1, (ii)の後半} \\ & = SM(\varphi,\ \omega_j) \quad \because \text{ axiom 2の(iii)} \end{aligned} \tag{2.55}$$

が得られるからである。式(2.55)の導出と同様にして、次の(ロ)の不変性も証明できる。

（ロ）(SMのU－不変性)

TのU－不変式(2.53)が成立していれば、

$$\forall \varphi \in \Phi,\ \forall j \in J,\ SM(U\varphi,\ \omega_j) = SM(\varphi,\ \omega_j) \tag{2.56}$$

3. support vector machineとしての、axiom 3を満たす大分類関数SMの設計と、その一般化

本章では，処理の対象とするパターン φ が選ばれた1つのカテゴリに帰属する可能性があるかどうかを決定できる“axiom 3を満たす大分類関数BSC”をsupport vector machine として構成する。

3.1 大分類関数 BSC

第j∈J番目のカテゴリ

$$\mathfrak{C}_j \in \underline{\mathfrak{C}} = \{\mathfrak{C}_j \mid j \in J\} \tag{3.1}$$

に帰属するパターン $\varphi \in \Phi$ については、

$$BSC(\varphi, j) = 1 \tag{3.2}$$

であり、かつ、第 $i \in J - \{j\}$ 番目のカテゴリ $\mathfrak{C}_i \in \underline{\mathfrak{C}}$ に帰属するパターン $\varphi \in \Phi$ については、

$$BSC(\varphi, j) = 0 \tag{3.3}$$

であるような機能を持つ、式(A.1)の大分類関数(binary-state classifier, rough classifier)BSC を，付録Aのaxiom 3を満たすように，hyperplane classifier としての形式

$$\sum_{\ell \in L} w(j, \ell) \cdot u(T\varphi, \ell) \tag{3.4}$$

の、0, 1への2値化変換

$$\begin{aligned} & BSC(\varphi, j) \\ & = psn\Big(\sum_{\ell \in L} w(j,\ \ell) \cdot u(T\varphi,\ \ell) + b(j)\Big) \end{aligned} \tag{3.5}$$

として、決定してみよう。ここに,

$$psn(u) = 0 \text{ if } u < 0,\ = 1 \text{ if } u \geqq 0 \tag{3.6}$$

であり、$\underline{w}(j) = \{w(j, \ell) \mid \ell \in L\}$ は the ℓ th weight vector であり，$-b(j)$ は the jth threshold である。また、式(1.6)の特徴抽出写像 u を導入し、パターン $\varphi \in \Phi$ から抽出される第 $\ell \in L$ 番目の特徴量を、$u(\varphi,\ \ell) \in R$(実数値全体の集合)とする。その全体を式(1.5)の $\underline{u}(\varphi)$ としよう。モデル構成作用素 T のべき等性(axiom 1の(iii)の後半)$TT = T$ を考慮すれば、axiom 3の(ii)(写像 T の下での**不変性**)が成立していることに注意しておく。

3.2 大分類関数 BSC の重みベクトル $\underline{w}(j)$ の決定に必要な最適化問題

We want to estimate a function BSC using input－output training data

$\langle \underline{u}(T\varphi_n),\ y_n\rangle \in R^{|L|}\times\{0, 1\}$ $(n=1\sim N)$

such that BSC will correctly classify unseen examples $\langle \underline{u}(T\varphi),\ y\rangle$, i.e. $BSC(\varphi, j)=y\in\{0, 1\}$. The example $\langle \underline{u}(T\varphi_n),\ y_n\rangle$ were generated from the same underlying probability distribution $P(\underline{u}(T\varphi), y)$ as the training data. □

the class of hyperplanes

$$HY_j: \sum_{\ell\in L} w(j, \ell)\cdot u(T\varphi, \ell)+b(j)=0,$$

where $\underline{w}(j)\in R^{|L|}$ and $b(j)\in R$ (3.7)

を考え、

$$z_n\cdot\left[\sum_{\ell\in L} w(j, \ell)\cdot u(T\varphi_n, \ell)+b(j)\right]\geqq 1,$$

where $z_n=2y_n-1=$

$$\begin{cases}+1 \text{ if } y_n=1\\ -1 \text{ if } y_n=0\end{cases} \tag{3.8}$$

が満たされるように、$\underline{w}(j)$を決めればよい。この決定問題を、the optimization problem としての最小化問題

minimize

$$2^{-1}\cdot\sum_{\ell\in L} w(j, \ell)^2, \tag{3.9}$$

subject to

$$z_n\cdot\left[\sum_{\ell\in L} w(j, \ell)\cdot u(T\varphi_n, \ell)+b(j)\right]\geqq 1$$

$$(n=1\sim N) \tag{3.10}$$

と考えてみよう。何故ならば、下記の補助定理3.1を勘案すれば、式(3.9)の値 $2^{-1}\cdot\sum_{\ell\in L} w(j, \ell)^2$ を最小化することは、$|L|$次元ユークリッド空間 $R^{|L|}$ 内の点 $P(=\underline{u}(T\varphi))$から式(3.12)で表される超平面 $HY_j=HY(\underline{w}(j),\ b(j))$に至る垂直距離

$$d(P, HY(\underline{w}(j),\ b(j))) \tag{3.11}$$

を最大化していることに相当しているからである。

次の補助定理3.1は幾何学ではよく知られている。

[補助定理3.1]

q次元ユークリッド空間 Rq 内の点P $\langle a_1, a_2, \cdots, a_q\rangle$ から超平面

$$HY: \sum_{\ell=1}^{q} w_\ell\cdot x_\ell+w_0=0 \tag{3.12}$$

に至る垂直距離 $d(P, HY)$は

$$d(P, HY)$$

$$=-\left[\sum_{\ell=1}^{q} w_\ell\cdot a_\ell+w_0\right]/\left[\varepsilon\cdot\sqrt{\Sigma_{\ell=1}^{q} w_\ell^2}\right] \tag{3.13}$$

と表される。ここに、

$$\varepsilon\in\{-1, +1\} \tag{3.14}$$

であり,

$$\sum_{\ell=1}^{q} w_\ell\cdot a_\ell+w_0\geqq 0\Rightarrow\varepsilon=-1$$

$$\sum_{\ell=1}^{q} w_\ell\cdot a_\ell+w_0<0\Rightarrow\varepsilon=+1 \tag{3.15}$$

と、選ぶ。 □

3.3 support vectors によるaxiom 3を満たす大分類関数 BSC の設計

3.3.1 Support Vector Machine SVM

SVMはC.Cortes, V.Vapnikによって提案された“重みベクトルが support vectors を陽に持つ1次結合によって与えられる2カテゴリ分類のための1次識別関数”である［A4］。

分類の対象となるものをパターンというが、S.Suzuki以外のパターン認識の理論は，パターンが既に圧縮されたものをパターンと称して論が展開されることが多い.既に圧縮されているものは事実上そのパターンの特徴量の組であるにもかかわらず。分類の前処理としてなされるデータ圧縮はこの意味で、ある程度似た者同士を1つにまとめるクラスタリングの働きをしている。

SVMの入力はパターンから抽出された特徴量の組である。

SVMとは、特徴量の組が低次元超平面によって線形分離可能でないとき，高次元超平面によって線形分離可能にすることを目的として設計される学習機械のことである。

それで、2つの｜L｜次元実数値ベクトル

$u = \mathrm{col}(u_1\ u_2 \cdots u_{|L|})$ (列ベクトル)

$$v = \mathrm{col}(v_1\ v_2 \cdots v_{|L|}) \tag{3.16}$$

間の内積［u, v］と、ノルム ｜u｜ を

$$[\underline{u}, \underline{v}] = \sum_{\ell \in L} u_\ell \cdot v_\ell \tag{3.17}$$

$$|\underline{u}| = [\sum_{\ell \in L} u_\ell{}^2]^{1/2} \tag{3.18}$$

と導入することになる。

Lagrange multipliers

$$\underline{\alpha} = \{\alpha_n > 0 \mid n = 1 \sim N\} \tag{3.19}$$

を導入し、a Lagrangian

$$f(\underline{w}(j), b(j), \alpha)$$
$$= 2^{-1} \cdot \sum_{\ell \in L} w(j, \ell)^2$$
$$- \sum_{n=1}^{N} \alpha_n \cdot \{z_n \cdot [\sum_{\ell \in L} \underline{w}(j, \ell) \cdot u(T\varphi_n, \ell) + b(j)] - 1\} \tag{3.20}$$
$$= 2^{-1} \cdot [\underline{w}(j), \underline{w}(j)] - [\underline{w}(j), \sum_{n=1}^{N} \alpha_n \cdot z_n \cdot$$
$$u(T\varphi_n)] - \{\sum_{n=1}^{N} \alpha_n \cdot z_n\} \cdot b(j)$$
$$+ \sum_{n=1}^{N} \alpha_n \tag{3.21}$$

が最小となる様に、w(j)，b(j)を決定する方程式系は、次の(イ)，(ロ)で与えられる：

$$(イ)\ \sum_{n=1}^{N} \alpha_n \cdot z_n = 0. \quad \therefore \quad \partial f / \partial b(j) = 0 \tag{3.22}$$

$$(ロ)\ w(j, \ell) = \sum_{n=1}^{N} \alpha_n \cdot z_n \cdot u(T\varphi_n, \ell),$$
$$\ell \in L. \quad \therefore \quad \partial f / \partial w(j, \ell) = 0, \ \ell \in L \tag{3.23}$$

□

2つの、互いに素な部分集合

$$[1, N]^+ \equiv \{n \in \{1, 2, \cdots, N\} \mid z_n = +1\} \tag{3.24}$$

$$[1, N]^- \equiv \{n \in \{1, 2, \cdots, N\} \mid z_n = -1\} \tag{3.25}$$

は共に空でないと見なすのが自然である：

$$\{1, 2, \cdots, N\} = [1, N]^+ \cup [1, N]^- \tag{3.26}$$

$$\wedge\ \phi\ (\text{the empty set}) = [1, N]^+ \cap [1, N]^-. \tag{3.27}$$

□

式(3.19)の$\underline{\alpha}$に関する制約条件式(3.22)は次の(3.29)のように表され、$[1, N]^+$, $[1, N]^-$に属する α_nの総和についての役割は同等であることがわかる.

[命題3.1]

$$\sum_{n=1}^{N} \alpha_n \cdot z_n = 0. \tag{3.28}$$

$$\Rightarrow$$

$$\Sigma_{n\in[1,N]^+}\alpha_n = \Sigma_{n\in[1,N]^-}\alpha_n. \tag{3.29}$$

(証明) $0 = \sum_{n=1}^{N} \alpha_n \cdot z_n$

$= \Sigma_{n\in[1,N]^+}\alpha_n \cdot z_n + \Sigma_{n\in[1,N]^-}\alpha_n \cdot z_n$

$= \Sigma_{n\in[1,N]^+}\alpha_n - \Sigma_{n\in[1,N]^-}\alpha_n.$ □

3.3.2 閾値 −b(j)の選び方

その後、閾値 $-b(j)$を、

$$-b(j) = 2^{-1}\cdot\left\{\max_{i\in J-\{j\}}[\underline{w}(j), \underline{u}(T\omega_i)] + [\underline{w}(j), \underline{u}(T\omega_j)]\right\} \tag{3.30}$$

と与えてみよう.このとき、

不等式

$$[\underline{w}(j), \underline{u}(T\omega_j)] - \max_{i\in J-\{j\}}[\underline{w}(j), \underline{u}(T\omega_i)] \geqq \beta(j) \geqq 0 \tag{3.31}$$

を満たす正数 $\beta(j)$が存在すると仮定すれば、

$$\sum_{\ell\in L}\underline{w}(j, \ell)\cdot\underline{u}(T\omega_j, \ell) + b(j)$$

$$= 2^{-1}\cdot\left\{[\underline{w}(j), \underline{u}(T\omega_j)] - \max_{i\in J-\{j\}}[\underline{w}(j), \underline{u}(T\omega_i)]\right\}$$

$$\geqq 2^{-1}\cdot\beta(j) \geqq 0 \tag{3.32}$$

$$\therefore$$

$$BSC(\omega_j, j) = 1 \quad \because \quad 2式(3.5), (3.6) \tag{3.33}$$

を得、都合がよい.

式(3.30)によって閾値 $-b(j)$を選ぶ方法を検討してみよう。不等式

$$\max_{i\in J-\{j\}}[\underline{w}(j), \underline{u}(T\omega_i)] < -b(j) \leqq [\underline{w}(j), \underline{u}(T\omega_j)] \tag{3.34}$$

を満たす実数 $b(j)$が存在するならば、不等式

$$[\underline{w}(j), \underline{u}(T\omega_j)] - \max_{i\in J-\{j\}}[\underline{w}(j), \underline{u}(T\omega_i)] > 0 \tag{3.35}$$

が成立し、不等式(3.31)を満たす正数$\beta(j)$が存在することがわかる。よって、例えば、閾値 $-b(j)$を式(3.30)の如く、選ぶことができる。

3.3.3 axiom 3を満たし、然も、各カテゴリ間の相互排除性大分類関数 BSC

constructing BSC from empirical data の立場から、 support vectors によるaxiom 3を満たす大分類関数 BSC を設計してみよう。

2式(3.33)は、axiom 3の(i)(カテゴリ抽出能力), であり、(ii)(T の下での不変性)の成立は

3.1節で示されているから、axiom 3を満たすように、式(A.1)のBSCが構成されたことになる。
更に、式(3.31)から、

$$
\begin{aligned}
&[\underline{w}(j),\ \underline{u}(T\omega_j)]+b(j)\\
&\quad-\max_{i\in J-\{j\}}[\underline{w}(j),\ \underline{u}(T\omega_i)]+b(j)】\\
&\geqq\beta(j)\geqq 0 \qquad (3.36)\\
&\quad\therefore\\
&[\underline{w}(j),\ \underline{u}(T\omega_j)]+b(j)\\
&\geqq\max_{i\in J-\{j\}}【[\underline{w}(j),\ \underline{u}(T\omega_i)]+b(j)】+\beta(j) \qquad (3.37)
\end{aligned}
$$

が成立しているけれども、

$$
\begin{aligned}
&0\geqq-\beta(j)\\
&>\max_{i\in J-\{j\}}【[\underline{w}(j),\ \underline{u}(T\omega_i)]+b(j)】 \qquad (3.38)\\
&\Rightarrow\\
&0>\max_{i\in J-\{j\}}【[\underline{w}(j),\ \underline{u}(T\omega_i)]+b(j)】+\beta(j) \qquad (3.39)\\
&\quad\therefore\ \ BSC(\omega_i,j)=0(i\neq j)\\
&\qquad\qquad\because\ \ 2式(3.5),\ (3.6) \qquad (3.40)
\end{aligned}
$$

を得る。式(%8)はカテゴリ間の相互排除性を表している。
結局、式(A.1)の大分類関数 BSCについて、2式(3.31), (3.38)が成立するような非負数$\beta(j)$が存在するような式(3.23)の重み$\underline{w}(j)$が等式(3.28)を満たしつつ、得られればよいことになる。
最適化問題のthe solution vector $\underline{w}(j)$は、$\alpha_n\neq 0$のとき、support vectors と称される$\underline{u}(T\varphi_n)$の観点から展開されているという。このとき、式(A.1)のBSCは support vector machine と呼ばれてよい。

3.4 ラグランジュ未定乗数の組 $\underline{\alpha}$ の、2次数理計画法による決定

式(3.19)の、正のLagrange multipliers $\underline{\alpha}$ を決定する方法を説明しよう。
式(3.23)の $\underline{w}(j,\ \ell)$から

$$w(j)=\sum_{n=1}^{N}\alpha_n\cdot z_n\cdot\underline{u}(T\varphi_n) \qquad (3.41)$$

と書けるから、この式(3.41)の$\underline{w}(j)$を式(3.21)の $f(\underline{w}(j),\ b(j),\ \underline{\alpha})$に代入すれば、

$$
\begin{aligned}
&f(w(j),\ b(j),\ \underline{\alpha})\\
&=2^{-1}\cdot[\underline{w}(j),\ \underline{w}(j)]-[\underline{w}(j),\ \underline{w}(j)]\\
&\quad-\{\sum_{n=1}^{N}\alpha_n\cdot z_n\}\cdot b(j)\\
&\qquad+\sum_{n=1}^{N}\alpha_n \qquad (3.42)\\
&=-2^{-1}\cdot[\underline{w}(j),\ \underline{w}(j)]\\
&\quad-\{\sum_{n=1}^{N}\alpha_n\cdot z_n\}\cdot b(j)\\
&\qquad+\sum_{n=1}^{N}\alpha_n \qquad (3.43)
\end{aligned}
$$

であるが、式(3.28)を考慮すれば、

$$
\begin{aligned}
&=-2^{-1}[\underline{w}(j),\ \underline{w}(j)]\\
&\quad+\sum_{n=1}^{N}\alpha_n \qquad (3.44)
\end{aligned}
$$

と簡単化され、これに式(3.41)の$\underline{w}(j)$を代入すれば、結局、

$f(\underline{w}(j),\ b(j),\ \underline{\alpha})$

$= -2^{-1}\cdot \sum_{m=1}^{N}\sum_{n=1}^{N} z_m \cdot z_n \cdot$

$[\underline{u}(T\varphi_m),\ \underline{u}(T\varphi_n)]\cdot \alpha_m \cdot \alpha_n + \sum_{n=1}^{N} \alpha_n$ (3.45)

と再表現される。よって、式(3.19)の $\underline{\alpha}$ は、制約条件

$\forall n \in \{1, 2, \cdots, N\},\ \ \alpha_n \geqq 0\ \ \wedge$ 式(3.28) (3.46)

の下で、式(3.45)の $f(\underline{w}(j), b(j), \underline{\alpha})$ を最大化する "2次計画問題の解" として得られる(双対問題)。

3.5 support vector machine としての、大分類関数 BSC の一般化

式(3.45)の $BSC(\varphi, j)$ は、

$BSC(\varphi, j)$

$= psn([\underline{w}(j),\ \underline{u}(T\varphi)] + b(j))$ (3.47)

と表現されることに留意すると、関数

$g: R^{|L|} \times R^{|L|} \rightarrow R$ (3.48)

を導入し、

$BSC(\varphi, j)$

$= psn(g(\underline{w}(j),\ \underline{u}(T\varphi)) + b(j))$ (3.49)

と一般化される。ここに、関数 g として、内積 $[\underline{x}, \underline{y}]$ を一般化すれば、次の3選定(一), (二), (三)が考えられてよい：

(一)(多項式)$g(\underline{x}, \underline{y})$

$= \{c(j)\cdot[\underline{x}, \underline{y}] + d(j)\}^k (k=1, 2, \cdots)$ (3.50)

ここに、$-c(j)\cdot|\underline{x}|\cdot|\underline{y}| + d(j) \geqq 0$ (3.51)

(二)(ガウス形関数)$g(\underline{x}, \underline{y})$

$= \exp[-(2\sigma^2)^{-1}\cdot|\underline{x} - \underline{y}|^2]$ (3.52)

ここに、$\sigma > 0$ (3.53)

(三)(シグモイド関数)$g(\underline{x}, \underline{y})$

$= \tanh(p(j)\cdot[\underline{x}, \underline{y}] - q(j))$ (3.54)

ここに、$-p(j)\cdot|\underline{x}|\cdot|\underline{y}| - q(j) \geqq 0$ (3.55)

□

4. むすび

画像内容を理解し、この内容を言語化する画像理解システムを構築することを目指した研究の前段階として、本論文は書かれた。

これまで、S.Suzukiは、可分な一般抽象ヒルベルト空間 [A1] ～ [A3] $\mathfrak{H}$ 上で稼働する2つの**情報システム**(万能性連想形パターン認識システム **RECOGNITRON** [B3], [B4], [B14], [B21]、パターンの系列を記憶し、それを想起的再生をする連想形記憶システム **MEMOTRON** [B2], [B11]、並びにマルチメディア処理用ファジィ・プロダクション・システム **FUZZITRON** [B23] を

提案し、その簡単な計算機シミュレーション［B7］～［B17］,［B20］を介し、その性能を確かめている。

機械(による)学習は、新しい情報技術(IT)の基幹の1つをなすテキスト分類(text classification)、パターン分類に応用されている［A18］。SVMでのマージンとは、サポートベクタの通る2超平面間の距離のことであるが、ランダム予測･分類より少量だけ良好な予測･分類が可能な弱学習器(weak learner)を組み合わせ、一層高度な分類器を設計できる手法(ブースティング；boosting)［B24］と同様なlarge margin classifierの1つとしての有用なSVMを、SS理論での大分類関数BSCとして使用するための研究がなされた。

S.Suzuki理論を適用し、外界を理解する能力を備えたシステムに必要なT, SM, BSCの内、新たに、パターン $\varphi \in \Phi$ から抽出された特徴量の組 $\underline{u}(\varphi)$ を入力とするSVMの構造を利用し、axiom 3を満たす式(1.4)の大分類関数 BSC を設計した。

SVM理論を素直に利用した成果しか得られていないが、十分実用に耐えるBSCが設計されたどうかは計算機シミュレーションして見る必要がある。

文　献　A

［A 1］吉田耕作, 河田敬義, 岩村つらね：“位相解析の基礎”, 岩波書店, May 1963

［A 2］青木利夫, 高橋渉：“集合･位相空間要論”, 培風館, Sept.1979

［A 3］Angus E.Taylor, David C.Lay：“Introduction to function analysis”, John Wiley & Sons, Inc., 1980

［A 4］edited Bernhard Schölkopf, Christopher J.C. Burges, Alexander J.Smolap：“Advances in Kernel Methods(Support Vector Learning)”, The MIT Press, Cambridge, 1998

［A 5］佐藤敦, 山田敬じ：“ニューラルネットによるパターン認識［V･完］－新しいニューラルネットモデル－”, 電子情報通信学会誌, Vol.83, No.1, pp.50-56, Jan.2000

［A 6］平博順, 春野雅彦：“Support Vector Machine によるテキスト分類における属性選択”, 情報処理学会論文誌, Vol.41, No.4, pp.1113-1123, Apr.2000

［A 7］Pierre M.L.Drezet, Robert F.Harrison：“A new method for sparsity control in support vector classification and regression”, Pattern Recognition, Vol.34, pp.111-125, 2001

［A 8］太原育夫：“認知情報処理(ニューロサイエンス＆テクノロジーシリーズ)”, オーム社, Mar.1991

［A 9］Gilbert G.Walter：“Wavelets and other orthogonal systems with applications”, CRC Press, Inc., 1994,

［A10］Andrzej Cichocki, Rolf Unbehauen：“Neural networks for optimization and signal processing”, John Wiley & Sons, Mar.1994

［A11］R.S.Michalski他編：“概念と規則の学習－例からの学習(知識獲得と学習シリーズ5)”, 共立出版, 電総研人工知能研究グループ他訳, May 1988

［A12］横井俊夫：“メディアを手掛かりとしたAI技術･研究の再構築”, 人工知能学会誌, vol.13, no.5, Sept.1998

［A13］中津良平：“人間の非論理情報をAIはどう取り扱うか？”, 人工知能学会誌, vol.14, no.2, Mar.1999

[A14] 長尾真他："マルチメディア情報学の基礎(岩波講座マルチメディア情報学１)", 岩波書店, Oct.1999
[A15] いや富仁, 萩原将文："ファジー推論ニューラルネットワークを用いた風景画像からの知識抽出と認識", 電子情報通信学会論文誌D-Ⅱ, vol.J82-D-Ⅱ, no.4, pp.685-693, Apr.1999
[A16] Richard L.Dykstra："An algorithm for restricted least squares regression", Journal of the American Statistical Association , vol.78, no.384, pp.837-842, Dec.1983
[A17] 多鹿秀継："情報処理過程としての人間の記憶(レクチャーシリーズ「認知科学」(第１回))", 人工知能学会誌, vol.16, no.1, pp.1111-1118, Jan.2001
[A18] 永田昌明, 平瀬順："テキスト分類－学習理論の「見本市」－(情報論的学習理論とその応用)", 情報処理(情報処理学会誌), vol.42, no.1, pp.32-37, Jan.2001
[A19] 鈴木基之, 牧野正三："隠れマルコフ網の構造に基づく話者間距離の計算法", 人工知能学会誌, vol.15, no.5, pp.871-877, Sept.2000
[A20] Corinna Cortes, Vladimir Vapnik："Support-vector networks", Machine Learning, vol.20, pp.273-297, 1995
[A21] Massimiliano Pontil, Alessandro Verri："Support vector machines for 3D object recognition", IEEE Trans. on Pattern Analysis and Machine Intelligence, vol.20, no.6, pp.637-646, June 1998
[A22] 前田英作："痛快！サポートベクトルマシーン－古くて新しいパターン認識手法－", 情報処理(情報処理学会誌), vol.42, no.7, pp.676-683, July 2001
[A23] Anastasios Tefas, Constantinue Kotropoulos, loannis Pitas："Using Support Vector Machines to Enhance the Performance of Elastic Graphh Matching for Frontal Face Authentication", IEEE Trans.on Pattern Analysis and Machine Intelligence, vol.23, no.7, pp.735-746, July 2001

文　献　B

[B 1] 鈴木昇一："認識工学", 柏書房, Feb.1975
[B 2] 鈴木昇一："ニューラルネットの新数理", 近代文芸社, Sept.1996
[B 3] 鈴木昇一："パターン認識問題の数理的一般解決", 近代文芸社, June 1997
[B 4] 鈴木昇一："認識知能情報論の新展開", 近代文芸社, Aug.1998
[B 5] 鈴木昇一："パターンのエントロピーモデル", 電子情報通信学会論文誌(D-Ⅱ), vol.J77-D-Ⅱ, no.10, pp.2220-2238, Nov.1994
[B 6] 鈴木昇一："パターン認識の数学的理論", 電子情報通信学会技術研究報告［パターン認識と学習, パターン認識と理解］, PRL84-6(第1部), PRL84-30, PRL84-38, PRL85-27, PRU86-8, PRU86-35, PRU86-69, PRU87-1, PRU87-28, PRU88-30, PRU89-1, PRU89-27, PRU89-40, PRU89-66, PRU89-77, PRU89-136, PRU90-5, PRU90-15, PRU90-29, PRU90-125, PRU91-1, PRU91-29, PRU91-42, PRU92-1, PRU92-18, PRU92-25, PRU92-89, PRU92-102(第28部), May 1984～Jan.1993
[B 7] 鈴木昇一："手書き漢字の側抑制効果的分解とその計算機シミュレーション", 情報処理学会誌, vol.15, no.12, pp.927-934, Dec.1974
[B 8] 鈴木昇一："画像情報量とその手書き漢字への応用", 画像電子学会誌, vol.4, no.1, pp.4-12,

Apr.1975
[B 9] 鈴木昇一："抽出された特徴による手書き漢字構造の再生"，情報処理学会誌, vol.18, no.11, pp.1115-1122, Nov.1977
[B10] 鈴木昇一："回転群と画像の分解･強調･構造化再構成に関する計算機シミュレーション"，情報研究(文教大学･情報学部)，no.4, pp.36-56, Dec.1983
[B11] 鈴木昇一："連想形記憶N器MEMOTRONと日本語単独母音系列の再生に関する計算機シミュレーション"，情報研究(文教大学･情報学部)，no.7, pp.14-29, Dec.1986
[B12] 鈴木昇一："多変量解析に基づく大分類関数の決定とその計算機シミュレーション"，情報研究(文教大学･情報学部)，no.10, pp.35-49, Dec.1989
[B13] 鈴木昇一："帰属係数法に基づく類似度、帰属関係あいまい度、認識情報量の計算機シミュレーション"，情報研究(文教大学･情報学部)，no.11, pp.51-68, Dec.1990
[B14] 鈴木昇一："構造受精法と日本語単独母音の認識"，情報研究(文教大学･情報学部)，no.18, pp.17-51, Dec.1998
[B15] 鈴木昇一，前田英明："有声破裂音の代表パターンの学習的決定と、その計算機シミュレーション"，情報研究(文教大学･情報学部)，no.20, pp.77-95, Dec.1998
[B16] 鈴木昇一，前田英明："変動エントロピーによる有声破裂音の順序付けと、その計算機シミュレーション"，情報研究(文教大学･情報学部)，no.21, pp.51-78, Mar.1999
[B17] 鈴木昇一："平均顔を用いた顔画像の2値化、並びに、目･鼻･口の抽出と、その計算機シミュレーション"，情報研究(文教大学･情報学部)，no.22, pp.65-150, Dec.1999
[B18] 鈴木昇一："直交系によるパターンモデルの構成"，情報研究(文教大学･情報学部)，no.21, pp.23-49, Mar.1999
[B19] 鈴木昇一："認識行為に向けての、効用最大化原理"，情報研究(文教大学･情報学部)，no.22, pp.151-210, Dec.1999
[B20] 鈴木昇一："界面エネルギーの減少に伴うモデル構成作用素の、顔画像処理に関する計算機シミュレーション"，情報研究(文教大学･情報学部)，no.23, pp.109-182, Mar.2000
[B21] 鈴木昇一："風景画から知識を抽出し、解釈するシステムの、ファジィ推論ニューラルネットによる構成"，情報研究(文教大学･情報学部)，no.23, pp.183-265, Mar.2000
[B22] 鈴木昇一："各個人の感性を反映した認識システムRECOGNITRON"，情報研究(文教大学･情報学部)，no. 24, pp.185-257, Dec.2000
[B23] 鈴木昇一："プロダクション･システムとしてのファジィ･マルチメディア･コンピュータと、空間多重パターンファジィ推論系"，情報研究(文教大学･情報学部)，no.24, pp.105-183, Dec.2000
[B24] 鈴木昇一："SS大分類関数BSCの適応的構成への、計算論的学習理論の適用"，情報研究(文教大学･情報学部)，no.25, pp.187-238, Mar.2001
[B25] 鈴木昇一："量子力学の諸原理，多段階量子認識系と，心理状態を取り入れた想起に基づく部分空間認識法"，情報研究(文教大学･情報学部)，no.25, pp.239-284, Mar.2001

付録A. axiom 3を満たす大分類関数 BSC

本付録Aでは、ある1つのカテゴリに帰属するどうかを決定する2カテゴリ分類器としての大分類関数 BSC はaxiom 3を満たすように構成されなければならないことを説明し、その後、構成例を掲げる。

A1. axiom 3と大分類関数 BSC

大分類関数(rough classifier, binary-state classifier)と呼ばれる**2値関数**

$$\mathrm{BSC} : \Phi \times \mathrm{J} \rightarrow \{0, 1\} \tag{A.1}$$

を、次のaxiom 3 を満たすものとして導入し、解釈

パターン $\varphi \in \Phi$ の帰属するカテゴリ候補の1つが
第 $j \in J$ 番目のカテゴリ $\mathfrak{C}_j$ であるならば、
$\mathrm{BSC}(\varphi, j) = 1$ であることが望ましい (A.2)

を採用しょう。この際、注意すべきは、

$\mathrm{BSC}(\varphi, j) = 0$ であっても、パターン $\varphi \in \Phi$ の
帰属するカテゴリ候補の1つは、第 $j \in J$ 番目の
カテゴリ $\mathfrak{C}_j$ でないとは限らない (A.3)

としていることである。また、axiom 3 の(i)からわかるように、**カテゴリ間の相互排除性**(the mutual exclusion of the one category from the other categories)

$$\forall j \in J, \forall i \in J - \{j\}, \ \mathrm{BSC}(\omega_i, j) = 0 \tag{A.4}$$

を公理として要請していない事実に注意しておこう。

<u>Axiom 3</u> (**大分類関数 BSC の満たすべき公理**)

(i)(**カテゴリ抽出能力**;category separability)

$\forall j \in J, \mathrm{BSC}(\omega_j, j) = 1.$

(ii)(写像 T の下での不変性;invariance under mapping T)

$\forall \varphi \in \Phi, \forall j \in J, \mathrm{BSC}(T\varphi, j) = \mathrm{BSC}(\varphi, j).$ □

BSC for the j-th category $\mathfrak{C}_j$ is trained to distinguish between patterns belonging to $\mathfrak{C}_j$ and its complement $\underline{\mathfrak{C}} - \{\mathfrak{C}_j\}$. In general, each category $\mathfrak{C}_j$ can have any number of exemplars. Even if there are roughly equal numbers of exemplars for each of the $|J|$ categories, $\mathfrak{C} - \{\mathfrak{C}_j\}$ will have many more exemplars than category $\mathfrak{C}_j$.

次の定理A.1は、大分類関数 BSC の出力 $\mathrm{BSC}(\varphi, j)$ がパターン変換 U に関し、不変に保たれるには、変換後のパターン $U\varphi$ が変換前のパターン φ と同一のパターンモデル $T\varphi$ を持てばよいことを明らかにしている。

[定理A.1] (**大分類関数 BSC のU－不変性**)

モデル構成作用素 T の U－不変式(2.51)が成立するような、パターン $\varphi \in \Phi$ と式(2.50)のパターン変換 U に関し、

$$\forall j \in J, \mathrm{BSC}(U\varphi, j) = \mathrm{BSC}(\varphi, j). \tag{A.5}$$

(証明) $\forall j \in J, \mathrm{BSC}(U\varphi, j)$

$= \mathrm{BSC}(T(U\varphi), j)$ ∵ axiom 3の(ii)

$= \mathrm{BSC}(T\varphi), j)$ ∵ 式(2.51)

$= \mathrm{BSC}(\varphi, j).$ ∵ axiom 3の(ii) □

A2. 大分類関数 BSC の構成例

本節では、axiom 3を満たし、然も、有用な式(A.1)の大分類関数 BSC を構成してみよう。

A2.1　包含情報量による axiom 2を満たす類似度関数 SM の構成

代表パターンモデル集合 $T\cdot\Omega$ についての非一致条件

$$\forall j\in J,\ \forall i\in J-\{j\},\ \| T\omega_i - T\omega_j \| > 0 \tag{A.6}$$

の下で、パターンモデル $T\varphi$ に含まれるパターンモデル $T\omega_j$ の量を情報量(amounut of information)として計量化すれば、

$$-2^{-1}\cdot\log_e\left[1-\left|(T\varphi\,\|T\varphi\|^{-1},\ T\omega_j\,\|T\omega_j\|^{-1})\right|^2\right] \tag{A.7}$$

であり、規格化すれば、

$$f_j(\varphi)=\begin{cases} -2^{-1}\cdot\log_e\left[1-\left|(T\varphi\,\|T\varphi\|^{-1},\ T\omega_j\,\|T\omega_j\|^{-1})\right|^2\right] \\ \quad /\sum_{k\in J} -2^{-1}\cdot\log_e\left[1-\left|(T\varphi\,\|T\varphi\|^{-1},\ T\omega_k\,\|T\omega_k\|^{-1})\right|^2\right. \\ \quad\cdots\exists i\in J,\ (T\varphi,\ T\omega_i)\neq 0 \\ -2^{-1}\cdot\log_e\left[1-p(\mathfrak{C}_j)\right]\ / \\ \quad /\sum_{k\in J} -2^{-1}\cdot\log_e\left[1-p(\mathfrak{C}_j)\right] \\ \quad\cdots\forall i\in J,\ (T\varphi,\ T\omega_i)=0 \end{cases} \tag{A.8}$$

で表される。包含情報量と称されてよいこの f_j を使用すれば、次の定理A.2の如く、axiom 2を満たす式(2.22)の関数 SM を構成できる。

［定理A.2］（**類似度関数 SM の、包含情報量 f_j による構成定理**）

$$SM(\varphi,\ \omega_j)=f_j(\varphi) \tag{A.9}$$

と定義された式(2.22)の関数 SM は、axiom 2を満たす。　□

A2.2　式(A.8)の $f_j(\varphi)$ の更新形式を採用した axiom 3を満たす類似度関数 SM の構成

さて、式(2.9)に登場している代表パターン集合 Ω に関する非正条件

$$\forall j\in J,\ \forall i\in J-\{j\},\ g_j(T\omega_i)\leqq 0 \tag{A.10}$$

を満たす関数 g_i の系

$$g_i : T\cdot\Phi\rightarrow R(\text{実数全体の集合}),\ i\in J \tag{A.11}$$

を用意した後、関数

$$hi : \Phi\rightarrow\{s\mid 0\leqq s\leqq 1\}\ (\text{単位区間の実数全体の集合}),\ i\in J \tag{A.12}$$

を、式(A.8)の $f_j(\varphi)$ が更新される形式で、次のように定義する。ここに、max{a, b} は2つの実数 a, b の内、小さくない方を指す：

$$h_j(\varphi)=\begin{cases} [f_j(\varphi)+\max\{g_j(T\varphi),\ 0\}] \\ \quad /\sum_{i\in J}[f_i(\varphi)+\max\{g_i(T\varphi),\ 0\}] \\ \quad\cdots\sum_{i\in J}[f_i(\varphi)+\max\{g_i(T\varphi),\ 0\}]=0\text{の場合} \\ p(\mathfrak{C}_j) \\ \quad\cdots\sum_{i\in J}[f_i(\varphi)+\max\{g_i(T\varphi),\ 0\}]=0\text{の場合} \end{cases} \tag{A.13}$$

□

このとき、次の4性質(イ),(ロ),(ハ),(ニ)が成立する：

(イ)(最大1性質) $\forall j \in J,\ h_j(\omega_j)=1.$

(ロ)(最小0性質)

$\forall j \in J,\ \forall i \in J-\{j\},\ h_j(\omega_i)=0.$

(ハ)(規格化条件) $\forall \varphi \in \Phi,\ \sum_{j \in J} h_j(\varphi)=1.$

(ニ)(写像 T の下での不変性)

$\forall \varphi \in \Phi,\ \forall j \in J,\ h_j(T\varphi)=h_j(\varphi).$ □

(イ)の成立は、定理A.2から

$$\forall J \in J,\ f_j(\omega_j)=1 \tag{A.14}$$

$$\forall j \in J,\ \forall i \in J-\{j\},\ f_j(\omega_i)=0 \tag{A.15}$$

が成立しているから、

$$\begin{aligned}
&\forall j \in J,\ h_j(\omega_j)= \\
&[f_j(\omega_j)+\max\{g_j(T\omega_j),\ 0\}] \\
&\quad / [\textstyle\sum_{i \in J} [f_i(\omega_j)+\max\{g_i(T\omega_j),\ 0\}] \\
&=[1+\max\{g_j(T\omega_j),\ 0\}] \\
&\quad / [1+\textstyle\sum_{i \in J} \max\{g_i(T\omega_j),\ 0\}] \\
&=[1+\max\{g_j(T\omega_j),\ 0\}] \\
&\quad / [1+\max\{g_j(T\omega_j),\ 0\}] \quad \because \text{式(A.10)} \\
&=1
\end{aligned} \tag{A.16}$$

を得、示された。

(ロ)の成立も、2式(A.14),(A.15)を使い、

$$\begin{aligned}
&\forall j \in J,\ \forall k \in J-\{j\}, \\
&h_j(\omega_k)= \\
&[f_j(\omega_k)+\max\{g_j(T\omega_k),\ 0\}] \\
&\quad / [\textstyle\sum_{i \in J} [f_i(\omega_k)+\max\{g_i(T\omega_k),\ 0\}] \\
&=\ [0+0]/[1+\max\{g_k(T\omega_k),\ 0\}] \\
&\qquad \because \text{式(A.10)} \\
&=0
\end{aligned} \tag{A.17}$$

を得、示された。

(ハ)の成立は、$h_j(\varphi)$の定義式(A.13)から明らかである。

最後の性質(ニ)は、axiom 1の(iii)の後半から、

$$\forall \varphi \in \Phi,\ \forall j \in J,\ f_j(T\varphi)=f_j(\varphi) \tag{A.18}$$

$$\wedge g_j(T(T\varphi))=g_j(\varphi) \tag{A.19}$$

が成立すること明らかである。

上述の4性質(イ),(ロ),(ハ),(ニ)の成立を表現し直せば、、次の定理A.3のようになる。

［定理A.3］(**類似度関数 SM の、関数 h_j による構成定理**)

$$SM(\varphi,\ \omega_j)=h_j(\varphi) \tag{A.20}$$

と定義された式(2.22)の関数 SM は、axiom 2を満たす。 □

A2.3 超平面によるaxiom 3を満たす大分類関数 BSC の構成

このとき、次の定理A.4によってaxiom 3を満たす式(A.1)の大分類関数 BSC が、その総和が 1となり 1 より大きくない非負実数値としての式(A.13)の $h_i(\varphi)$の組が式(A.21)の実数値重み W(j, i)の組から定まる超平面の、正負のどちら側にあるかを判定することにより、定まることがわかる。

[定理A.4] (**大分類関数 BSC の構成定理**)

実数値重み W(j, i)の組

$$W(j,i),\ \ j\in J,\ i\in J\cup\{0\} \tag{A.21}$$

に関する非負条件

$$\forall j\in J,\ W(j,j)+W(j,0)\geqq 0 \tag{A.22}$$

の下で、

$$\mathrm{BSC}(\varphi,j)=\begin{cases}1\cdots\sum_{i\in J}W(j,i)\cdot h_i(\varphi)+W(j,0)\geqq 0 \text{ のとき}\\ 0\cdots\sum_{i\in J}W(j,i)\cdot h_i(\varphi)+W(j,0)<0 \text{ のとき}\end{cases} \tag{A.23}$$

と定義された式(A.1)の関数 BSC は、axiom 3を満たす。更に、負条件

$$\forall j\in J,\ \forall k\in J-\{j\},\ \ W(j,k)+W(j,0)<0 \tag{A.24}$$

の下で、カテゴリ間の相互分離条件

$$\forall j\in J,\ \forall k\in J-\{j\},\ \ \mathrm{BSC}(\omega_k,j)=0 \tag{A.25}$$

も成立する。

(証明) A2.2節の(イ), (ロ)を考慮すれば、axiom 3, (i)の成立は、

$$\begin{aligned}&\forall j\in J,\\ &0\leqq\sum_{i\in J}W(j,i)\cdot h_i(\omega_j)+W(j,0)\\ &=W(j,j)+W(j,0)\\ &\Rightarrow \mathrm{BSC}(\omega_j,j)=1\end{aligned} \tag{A.26}$$

を得、示された。

axiom 3, (ii)の成立は、A2.2節の(ニ)から明らかである。

カテゴリ間の相互分離条件式(A.25)の成立は、A2.2節の(イ), (ロ)を考慮すれば、

$$\begin{aligned}&\forall j\in J,\ \forall k\in J-\{j\},\\ &0>\sum_{i\in J}W(j,i)\cdot h_i(\omega_k)+W(j,0)\\ &=W(j,k)+W(j,0)\\ &\Rightarrow \mathrm{BSC}(\omega_k)=0\end{aligned}$$

を得、示された。 □

2式(A.22), (A.24)をまとめると、

$$\begin{aligned}&\forall j\in J,\ \forall k\in J-\{j\},\\ &W(j,j)\geqq -W(j,0)>W(j,k)\end{aligned} \tag{A.27}$$

ということになる。

A2.4 式(A.13)の $h_j(\varphi)$内に登場している式(A.11)の $g_j(\varphi)$の、特徴抽出写像 u を用いた1次ニューラルネットによる選定法

非正条件式(A.10)を満たす式(A.11)の g_j は式(A.13)の $h_j(\varphi)$内に登場しているが、実数重み $V_1(j,\ell)$の組、実数閾値 $V_0(j)$の組

$$V_1(j,\ell),\ \ \ell\in L,\ V_0(j)\ (j\in J) \tag{A.28}$$

を用いて、

$$g_j(T\varphi) = \sum_{\ell\in L} V_1(j,\ \ell)\cdot u(T\varphi,\ \ell)+V_0(j) \tag{A.29}$$

と、設定することが考えられる。ここに、

$$u: \Phi\times L\rightarrow R(\text{実数全体の集合}) \tag{A.30}$$

は特徴抽出写像であり、$u(\varphi,\ \ell)\in R$ は、

パターン $\varphi\in\Phi$ から抽出された第 $\ell\in L$ 番目の特徴量 (A.31)

である。

$$\forall j\in J,\ \forall i\in J-\{j\},$$
$$g_j(T\omega_i)= \sum_{\ell\in L} V_1(j,\ \ell)\cdot u(T\omega_i,\ \ell)+V_0(j)\leqq 0 \tag{A.32}$$

を満たすように、各 V_1, V_2 が学習で決定されていれば、非正条件式(A.10)が満たされることがわかる。

分離がよくなるためには、

$$\forall j\in J,\ g_j(T\omega_j)>0 \tag{A.33}$$

つまり、

$$\forall j\in J,\ g_j(T\omega_j)= \sum_{\ell\in L} V_1(j,\ \ell)\cdot u(T\omega_j,\ \ell)+V_0(j)>0 \tag{A.34}$$

であることが望ましい。

A2.5 式(A.13)の $h_j(\varphi)$内に登場している式(A.8)の $f_j(\varphi)$の、今1つの選定法

式(A.13)の $h_j(\varphi)$内に登場している式(A.8)の$f_j(\varphi)$の選定法には、任意性がある。別の選び方を考えよう。

A2.5.1 相違度関数 dsm_j の5種類構成

$|J|$ 個のパターン集合 Ψ_j の系

$$\Psi_j(\subset\Phi),\ j\in J \tag{A.35}$$

を考え、包含条件

$$\forall j\in J,\ T\cdot\omega_j\in T\cdot\Psi_j \tag{A.36}$$

と、非一致条件

$$\forall i\in J,\ \forall j\in J,\ \psi\in\Psi i,\ \psi'\in\Psi_j,$$
$$\|T\psi-T\psi'\|>0 \tag{A.37}$$

とを満たしているとしよう。

2つのパターンが似ていないほど大きい値をとるような非類似度関数(dissimilarity measure function) dsmj の系

$$dsm_j: \Phi\rightarrow R^+(\text{非負実数全体の集合}),\ j\in J \tag{A.38}$$

が、3性質(一), (二), (三)を満たすように選ばれているとしよう：

(一)(最小0性質) $\forall j\in J,\ \forall\psi\in\Psi_j,\ dsm_j(\psi)=0.$

(二)(正性質)

$\forall j\in J,\ \forall i\in J-\{j\},\ \ \psi\in\Psi_i,\ dsm_j(\psi)>0.$

(三)(写像 T の下での不変性)

$\forall\varphi\in\Phi,\ \forall j\in J,\ dsm_j(T\varphi)=dsm_j(\varphi).$

□

例えば、次の典型的な5種類の $\mathrm{dsm}_j(\varphi)$ は、次の3性質(一)，(二)，(三)を満たすことがわかる。特に、(三)はaxiom 1の(iii)の後半から成り立つ：

①$\mathrm{dsm}_j(\varphi)=\min_{\psi\in\Psi_j}\ \|T\varphi-T\psi\|.$ (A.39)

②$\mathrm{dsm}_j(\varphi)=\min_{\psi\in\Psi_j}\ \log_e[1+\|T\varphi-T\psi\|^{|J|}\}.$ (A.40)

③$\mathrm{dsm}_j(\varphi)=\min_{\psi\in\Psi_j}\ [1-\exp[-a_j^{-1}\cdot\|T\varphi-T\psi\|^2]].$ (A.41)

ここに、a_j は正定数であり、

$a^j=$

$3^{-1}\cdot\min_{\psi\in\Psi_i}\ \min_{\psi'\in\Psi_j(T\varphi\neq T\psi')}\|T\psi-T\psi'\|^2$ (A.42)

④$\mathrm{dsm}_j(\varphi)=\min_{\psi\in\Psi_j}[1-|\mathrm{nip}(T\varphi,T\psi)|^2].$ (A.43)

ここに、$\mathrm{nip}(T\varphi, T\psi)$ は2つのパターンモデル $T\varphi, T\psi$ の規格化内積(normalized inner product)であり、

$$\mathrm{nip}(T\varphi,T\psi)\equiv\begin{cases}(T\varphi,T\psi)/[\|T\varphi\|\cdot\|T\psi\|] \\ \quad\cdots\|T\varphi\|\cdot\|T\psi\|>0\text{ のとき}\\ 0\cdots\|T\varphi\|\cdot\|T\psi\|=0\text{ のとき.}\end{cases}\tag{A.44}$$

⑤$\mathrm{dsm}_j(\varphi)=$

$$\begin{cases}\min_{\psi\in\Psi_j}-2^{-1}\cdot\log_e|\mathrm{nip}(T\varphi,T\psi)|^2\\ \quad\cdots[\forall\psi\in\Psi_j,\ \|T\varphi\|\cdot\|T\psi\|>0]\\ \qquad\wedge|\mathrm{nip}(T\varphi,T\psi)|^2>\varepsilon_j\text{ のとき}\\ -2^{-1}\cdot\log_e\varepsilon_j\\ \quad\cdots[\exists\psi\in\Psi_j,\ \|T\varphi\|\cdot\|T\psi\|=0]\\ \qquad\vee|\mathrm{nip}(T\varphi,T\psi)|^2\leqq\varepsilon_j\text{ のとき.}\end{cases}\tag{A.45}$$

ここに、正定数 ε_j は、

$0<\varepsilon_j<\min_{i\in J-\{j\}}\ \min_{\psi\in\Psi_j}\ \min_{\psi'\in\Psi_i}|\mathrm{nip}(T\psi,T\psi')|^2$ (A.46)

□

A2.5.2 非規格化類似度関数 $\mathrm{sim}_j(\varphi)$ の構成

前項の3性質(一)，(二)，(三)を満たす式(A.38)の非類似度関数 dsm_j を使って、関数 sim_j の系

$\mathrm{sim}_j:\Phi\to R^+,\ j\in J$ (A.47)

を

$\mathrm{sim}_j(\varphi)=\min_{i\in J-\{j\}}\ \mathrm{dsm}_i(\varphi)$ (A.48)

と定義すれば、各 simj(φ)は、次の3性質(一)，(二)，(三)を満た、非規格化類似度関数と解釈されてよい：

(一)(正性質)$\forall j\in J,\ \forall\psi\in\Psi_j,\ \mathrm{sim}_j(\psi)>0.$

(二)(最小0性質)

$\forall j\in J,\ \forall i\in J-\{j\},\ \psi\in\Psi_i,\ \mathrm{sim}_j(\psi)=0.$

(三)(写像 T の下での不変性)

$\forall\varphi\in\Phi,\ \forall j\in J,\ \mathrm{sim}_j(T\varphi)=\mathrm{sim}_j(\varphi).$ □

A2.5.3 axiom 2を満たす類似度関数 $f_j(\varphi)$ の構成

前項の3性質(一)，(二)，(三)を満たす式(A.48)の類似度関数 sim_j を規格化して、

$$
f_j(\varphi)=
\begin{cases}
\mathrm{sim}_j(\varphi)/\sum_{k\in J}\mathrm{sim}_k(\varphi) \\
\qquad \cdots \sum_{k\in J}\mathrm{sim}_k(\varphi)>0 \text{ のとき} \\
p(\mathfrak{C}_j)\cdots \\
\qquad \cdots \sum_{k\in J}\mathrm{sim}_k(\varphi)=0 \text{ のとき}
\end{cases}
\tag{A.49}
$$

と定義される関数 f_j の系

$$
f_j : \Phi \to \{s \mid 0\leqq s\leqq 1\},\ j\in J \tag{A.50}
$$

を定義しよう。

この f_j を使用すれば、次の定理A.5の如く、axiom 2を満たす式(2.22)の関数 SM を構成できる。

［定理A.5］(**類似度関数 SM の、f_j による構成定理**)

$$
\mathrm{SM}(\varphi,\ \omega_j)=f_j(\varphi) \tag{A.51}
$$

と定義された式(2.22)の関数 SM は、axiom 2を満たす。

(証明)$T\omega_j$ の包含式(A.36)を考慮すれば、以下の(ア),(イ)の証明から、axiom 2,(i)が成立することがわかる。

(ア)$\forall j\in J,\ \forall \psi\in\Psi_j$,

$$
\begin{aligned}
\mathrm{SM}(\psi,\ \omega_j)&=f_j(\psi) \quad \because \text{ 式(A.51)} \\
&=\mathrm{sim}_j(\psi)/\ [\mathrm{sim}_j(\psi)+\sum_{k\in J-\{j\}}\mathrm{sim}_k(\psi)] \\
&\qquad \because \text{ 式(A.49)} \\
&=0 \quad \because \text{ 2.5.2項の2性質(一),(二).}
\end{aligned}
\tag{A.52}
$$

(イ)$\forall j\in J,\ \forall i\in J-\{j\},\ \forall \psi\in\Psi_i$,

$$
\begin{aligned}
\mathrm{SM}(\psi,\ \omega_j)&=f_j(\psi) \quad \because \text{ 式(A.51)} \\
&=\mathrm{sim}_j(\psi)/\ [\mathrm{sim}_i(\psi)+\sum_{k\in J-\{i\}}\mathrm{sim}_k(\psi)] \\
&\qquad \because \text{ 式(A.49)} \\
&=0/\ [\mathrm{sim}_i(\psi)+\sum_{k\in J-\{i\}}0] \\
&\qquad \because \text{ 2.5.2項の性質(二)} \\
&=0 \quad \because \text{ 2.5.2項の性質(一).}
\end{aligned}
\tag{A.53}
$$

axiom 2,(ii)(規格化性)の成立は、f_j の定義式(A.49)から明らかである。

axiom 2,(iii)(T の下での不変性)の成立は、2.5.2項の、各 sim_j の性質(三)から明らかである。

□

A2.6 2次ニューラルネットによるaxiom 3を満たす大分類関数 BSC の構成

式(A.23)による大分類関数BSC の設定は、1次ニューラルネット、いわゆるパーセプトロン(perceptron)の構造形式を利用したものである。本節では、2次ニューラルネットの構造形式を利用して、

$$
\begin{aligned}
&\mathrm{BSC}(\varphi, j) \\
&=\mathrm{psn}\Big(\sum_{i\in J}\sum_{k\in J} W(j, i, k)\cdot h_i(\varphi)\cdot h_k(\varphi) \\
&\quad +\sum_{i\in J} W(j, i)\cdot h_i(\varphi)+W(j, 0)\Big)
\end{aligned}
\tag{A.54}
$$

と設定してみよう。ここに、1実変数 u の2値関数 psn は、

$$
\mathrm{psn}(u)=0 \text{ if } u<0,\ =1 \text{ if } u\geqq 0 \tag{A.55}
$$

と定義される。

このとき、次の定理A.6によってaxiom 3を満たす式(A.1)の大分類関数BSCが、その総和が1となり1より大きくない非負実数値としての式(A.13)の$h_i(\varphi)$の組が実数値重み$W(j, i, k)$，$W(j, i)$の組から定まる曲面の、正負のどちら側にあるかを判定することにより、定まることがわかる。

［定理A.6］**(2次ニューラルネットによる、大分類関数 BSC の構成定理)**

実数値重み$W(j, i, k)$，$W(j, i)$の組

$$W(j, i, k),\ j, i, k \in J$$

$$W(j, i),\ j \in J, i \in J \cup \{0\} \tag{A.56}$$

に関する非負条件

$$\forall j, i \in J,\ W(j, j, j) + W(j, j) + W(j, 0) \geqq 0 \tag{A.57}$$

の下で、式(A.54)の如く定義された式(A.1)の関数BSCは、axiom 3を満たす。更に、負条件

$$\forall j \in J,\ \forall k \in J - \{j\},$$

$$W(j, k, k) + W(j, k) + W(j, 0) < 0 \tag{A.58}$$

の下で、カテゴリ間の相互分離条件

$$\forall j \in J,\ \forall k \in J - \{j\},\ BSC(\omega_k, j) = 0 \tag{A.59}$$

も成立する。

(証明)　A2.2節の(イ)，(ロ)を考慮すれば、axiom 3，(ⅰ)の成立は、

$$\forall j \in J,$$

$$\begin{aligned}
&0 \leqq \sum_{i \in J} \sum_{k \in J} W(j, i, k) \cdot h_i(\omega j) \cdot h_k(\omega_j) \\
&+ \sum_{i \in J} W(j, i) \cdot h_i(\omega_j) + W(j, 0) \\
&= W(j, j, j) + W(j, j) + W(j, 0) \\
&\Rightarrow BSC(\omega_j, j) = 1
\end{aligned} \tag{A.60}$$

を得、示された。

axiom 3，(ⅱ)の成立は、A2.2節の(ニ)から明らかである。

カテゴリ間の相互分離条件式(A.59)の成立は、A2.2節の(イ)，(ロ)を考慮すれば、

$$\forall j \in J,\ \forall \ell \in J - \{j\},$$

$$\begin{aligned}
&0 > \sum_{i \in J} \sum_{k \in J} W(j, i, k) \cdot h_i(\omega_\ell) \cdot h_k(\omega_\ell) \\
&+ \sum_{i \in J} W(j, i) \cdot h_i(\omega_\ell) + W(j, 0) \\
&= W(j, \ell, \ell) + W(j, \ell) + W(j, 0) \\
&\Rightarrow BSC(\omega_\ell) = 0
\end{aligned} \tag{A.61}$$

を得、示された。　□

2式(A.57)，(A.58)をまとめると、

$$\forall j \in J,\ \forall k \in J - \{j\},$$

$$\begin{aligned}
&W(j, j, j) + W(j, j) \geqq -W(j, 0) \\
&> W(j, k, k) + W(j, k)
\end{aligned} \tag{A.62}$$

ということになる。

A3. 式(A.13)の$h_j(\varphi)$内に登場している式(A.11)の$g_j(\varphi)$の、特徴抽出写像uを用いた2次ニューラルネットによる選定法

A3.1　$g_j(\varphi)$の、2次ニューラルネットによる選定

非正条件式(A.10)を満たす式(A.11)のg_jは式(A.13)の$h_j(\varphi)$内に登場しているが、A2.4節では、

式(A.11)の $g_j(\varphi)$ を式(A.29)の如く、パターンモデル $T\varphi$ から抽出された各特徴量 $u(T\varphi, \ell)$ を入力する1次ニューラルネットによって選定する手法が説明された。本章では、パターンモデル $T\varphi$ から抽出された各特徴量 $u(T\varphi, \ell)$ を入力するような、実数重み $V_2(j, k, \ell)$ の組、実数重み $V_1(j, \ell)$ の組、実数閾値 $V_0(j)$ の組

$$V_2(j, k, \ell),\ k,\ \ell \in L \tag{A.63}$$
$$V_1(j, \ell),\ \ell \in L \tag{A.64}$$
$$V_0(j) \tag{A.65}$$
$$, j \in J$$

を持つ2次ニューラルネットによって、各 $g_j(T\varphi)$ を、

$$\begin{aligned} &g_j(T\varphi) \\ &= \sum_{k\in L}\sum_{\ell\in L} V_2(j, k, \ell)\cdot u(T\varphi, k)\cdot u(T\varphi, \ell) \\ &+ \sum_{\ell\in L} V_1(j, \ell)\cdot u(T\varphi, \ell) + V_0(j) \end{aligned} \tag{A.66}$$

と、選定してみよう。

$$\begin{aligned} &\forall j \in J,\ \forall i \in J-\{j\}, \\ &g_j(T\omega_i) = \\ &\sum_{k\in L}\sum_{\ell\in L} V_2(j, k, \ell)\cdot u(T\omega_i, k)\cdot u(T\omega_i, \ell) \\ &+ \sum_{\ell\in L} V(j, \ell)\cdot u(T\omega_i, \ell) + V_0(j) \leqq 0 \end{aligned} \tag{A.67}$$

が成立するように、各 V_2, V_1, V_0 が学習で決定されていれば、非正条件式(A.10)が満たされることがわかる。

分離がよくなるためには、

$$\forall j \in J,\ g_j(T\omega_j) > 0 \tag{A.68}$$

つまり、

$$\begin{aligned} &\forall j \in J,\ g_j(T\omega_j) = \\ &\sum_{k\in L}\sum_{\ell\in L} V_2(j, k, \ell)\cdot u(T\omega_j, k)\cdot u(T\omega_j, \ell) \\ &+ \sum_{\ell\in L} V(j, \ell)\cdot u(T\omega_j, \ell) + V(j, 0) > 0 \end{aligned} \tag{A.69}$$

であることが望ましい。

A3.2 各 V_2, V_1, V_0 の学習による決定法

認識システムRECOGNITRONがパターン $\varphi \in \Phi$ に関し持っているカテゴリ帰属知識

$$\langle \varphi, \gamma \rangle \in \langle \Phi, 2^J \rangle$$

とは、パターンが φ カテゴリ部分集合 $\mathfrak{C}_j$, $j \in \gamma \in 2^J$ のいずれか1つのカテゴリに帰属している可能性があることをいう [B3]、[B4]。

2不等式(A.67), (A.69)を満たすように、訓練パターンのカテゴリ帰属知識系列

$$\langle \eta_0, [j_0] \rangle, \langle \eta_1, [j_1] \rangle, \cdots, \langle \eta_t, [j_t] \rangle, \cdots \tag{A.70}$$

を用いて、各 V_2, V_1, V_0 を学習の働きで決定する手法を以下で説明しよう。ここに、訓練パターン η_t は第 j_t 番目のカテゴリ $\mathfrak{C}_{j_t}$ に帰属していることが判明しているとしている。各カテゴリ $\mathfrak{C}_{j_t}$ に帰属する訓練パターン η_t はその生起確率 $p(\mathfrak{C}_{j_t})$ に比例する割合で生起しているものとし、各代表パターン $\omega_j (j \in J)$ は、

$$\forall j \in J,\ \exists t,\ \eta_t = \omega_j \tag{A.71}$$

というように、式(A.70)の系列に含まれているとしておかねばならない。

ある正定数$c_j>0(j\in J)$をあらかじめ、選定しておいて、

$$g_j(T\eta_t)=\begin{cases} c_j & \text{if } j=jt \\ -c_j & \text{if } j\in J-\{j_t\} \end{cases} \quad (A.72)$$

を満たすように、各V_2, V_1, V_0を逐次的に決定していけば、2不等式(A.67), (A.69)が満たされることになる。

式(A.70)の訓練系列は連続時刻$t(\geqq 0)$で与えられていると考えて、最急降下法を適用することを考えよう。

$$\mathrm{sgn}(j, j_t)=1 \text{ if } j\neq j_t, \ =-1 \text{ if } j=j_t \quad (A.73)$$

と定義される符号関数$\mathrm{sgn}(j, j_t)$を導入し、汎関数

$$F(\langle\eta_t, [j_t]\rangle) \equiv \sum_{i\in J} 2^{-1}\cdot[g_i(T\eta_t)-s(i, j_t)\cdot c_i]^2 \quad (A.74)$$

を定義し、3初期条件

$$V_2(j, k, \ell ; t) \mid_{t=0}=[\,|J|+2\cdot|L|+1]^{-1} \quad (A.75)$$

$$V_1(j, \ell ; t) \mid_{t=0}=[\,|J|+|L|+1]^{-1} \quad (A.76)$$

$$V_0(j ; t) \mid_{t=0}=[\,|J|+1]^{-1} \quad (A.77)$$

$$, j\in J, k\in L, \ell\in L$$

の下で、3学習方程式

①$$dV_2(j, k, \ell ; t)/dt =-\varepsilon_2(j, k, \ell ; t)\cdot \partial F(\langle\eta_t, [j_t]\rangle)/\partial V_2(j, k, \ell ; t) \quad (A.78)$$

②$$dV_1(j, \ell ; t)/d_t=-\varepsilon_1(j, \ell ; t)\cdot \partial F(\langle\eta_t, [j_t]\rangle)/\partial V_1(j, \ell ; t) \quad (A.79)$$

③$$dV_0(j ; t)/d_t=-\varepsilon_0(j ; t)\cdot \partial F(\langle\eta_t, [j_t]\rangle)/\partial V_0(j ; t) \quad (A.80)$$

の解として、各V_2, V_1, V_0を

$$V_2(j, k, \ell)=\lim_{t\to\infty} V_2(j, k, \ell ; t) \quad (A.81)$$

$$V_1(j, \ell)=\lim_{t\to\infty} V_1(j, \ell ; t) \quad (A.82)$$

$$V_0(j)=\lim_{t\to\infty} V_0(j ; t) \quad (A.82)$$

と求めればよい。ここに、$\varepsilon_2(j, k, \ell ; t)$, $\varepsilon_1(j, \ell ; t)$, $\varepsilon_0(j ; t)$は

$$\lim_{t\to\infty} \varepsilon_2(j, k, \ell ; t)=\lim_{t\to\infty} \varepsilon_1(j, \ell ; t) =\lim_{t\to\infty} \varepsilon_0(j ; t)=0 \quad (A.83)$$

を満たす正値非増加関数である。何故ならば、$V_2(j, k, \ell)$については、汎関数$F(\langle\eta_t, [j_t]\rangle)$の訓練時刻変数$t$についての単調非増加性

$$dF(\langle\eta_t, [j_t]\rangle)/dt =\sum_{j\in J}\sum_{k\in L}\sum_{\ell\in L} \partial F(\langle\eta_t, [j_t]\rangle)/ \partial V_2(j, k, \ell ; t)\cdot dV_2(j, k, \ell ; t)/d_t =-\sum_{j\in J}\sum_{k\in L}\sum_{\ell\in L} \varepsilon_2(j, k, \ell ; t)\cdot [\partial F(\langle\eta_t, [j_t]\rangle)/\partial V_2(j, k, \ell ; t)]^2$$

$\because$ 式(A.78)

$\leqq 0 \quad \because \quad \varepsilon_2(j, k, \ell ; t)$ は正値関数 (A.84)

が成立しているからである。$V_1(j, \ell)$, $V_0(j)$ についても同様である。

実際には、式(A.70)の訓練系列は離散時刻 $t=0, 1, 2, \cdots$ で与えられているから、3式(A.78)～(A.79)の、離散時刻表現を求めておかねばならない。

$$\varepsilon_2'(j, k, \ell ; t),\quad \varepsilon_1'(j, \ell ; t),\quad \varepsilon_0'(j ; t) \tag{A.85}$$

を正値関数として、例えば、

$$\varepsilon_2'(j, k, \ell ; t) = [j+k+\ell+t]^{-1} \tag{A.86}$$

$$\varepsilon_1'(j, \ell ; t) = [j+\ell+t]^{-1} \tag{A.87}$$

$$\varepsilon_0'(j ; t) = [j+t]^{-1} \tag{A.88}$$

と与えると、

④$V_2(j, k, \ell ; t+1)$

$$= V_2(j, k, \ell ; t) + \Delta V_2(j, k, \ell ; t) \tag{A.89}$$

, where

$\Delta V_2(j, k, \ell ; t)$

$$= -\varepsilon_2'(j, k, \ell ; t) \cdot \partial F(\langle \eta_t, [j_t] \rangle) / \partial V_2(j, k, \ell ; t) \tag{A.90}$$

⑤$V_1(j, \ell ; t+1)$

$$= V_1(j, \ell ; t) + \Delta V_1(j, \ell ; t) \tag{A.91}$$

, where

$\Delta V_1(j, \ell ; t)$

$$= -\varepsilon_1'(j, k, \ell ; t) \cdot \partial F(\langle \eta_t, [j_t] \rangle) / \partial V_1(j, \ell ; t) \tag{A.92}$$

⑥$V_0(j ; t+1)$

$$= V_0(j ; t) + \Delta V0(j ; t) \tag{A.93}$$

, where

$\Delta V_0(j ; t)$

$$= -\varepsilon_0'(j ; t) \cdot \partial F(\langle \eta_t, [j_t] \rangle) / \partial V0(j ; t) \tag{A.94}$$

$t=0, 1, 2, \cdots$

が求める離散時刻表現である。3式(A.90), (A.92), (A.94)に登場している偏微分係数 $\partial F(\langle \eta_t, [j_t] \rangle) / \partial V_2(j, k, \ell ; t)$, $\partial F(\langle \eta_t, [j_t] \rangle) / \partial V_1(j, \ell ; t)$, $\partial F(\langle \eta_t, [j_t] \rangle) / \partial V_0(j ; t)$ は、式(A.66)の $g_j(T\varphi)$ と、式(A.74)の $F(\langle \eta_t, [j_t] \rangle)$ とから、次のように求まる：

⑦$\partial F(\langle \eta_t, [j_t] \rangle) / \partial V_2(j, k, \ell ; t)$

$= \sum_{i \in J} [g_i(T\eta_t) - s(i, j_t) \cdot c_i] \cdot$

$$u(T\eta_t, k) \cdot u(T\eta_t, \ell) \tag{A.95}$$

⑧$\partial F(\langle \eta_t, [j_t] \rangle) / \partial V_1(j, \ell ; t)$

$$= \sum_{i \in J} [g_i(T\eta_t) - s(i, j_t) \cdot c_i] \cdot u(T\eta_t, \ell) \tag{A.96}$$

⑨$\partial F(\langle \eta_t, [j_t] \rangle) / \partial V_0(j ; t)$

$$= \sum_{i \in J} [g_i(T\eta_t) - s(i, j_t) \cdot c_i] \tag{A.97}$$

□

A4. パターンモデル Tφ を不変に保つパターン変換 U からもたらされる大分類関数 BSC の不変性

モデル構成作用素 T があるパターン変換 U に対し不変ならば、大分類関数 BSC もパターン変

換 U に対し不変であることは、次の(ロ)からわかる。

(イ)(BSCの正定数倍不変性)任意の正定数 a について、

$$\forall \varphi \in \Phi,\ \forall j \in J,$$

$$BSC(a \cdot \varphi,\ \omega_j) = BSC(\varphi,\ \omega_j) \tag{A.98}$$

が成立する。何故ならば、

$$\begin{aligned} &BSC(a \cdot \varphi,\ \omega_j) \\ &= BSC(T(a \cdot \varphi),\ \omega_j) \quad \because\ \text{axiom 3の(ii)} \\ &= BSC(T\varphi,\ \omega_j) \quad \because\ \text{axiom 1, (ii)の後半} \\ &= BSC(\varphi,\ \omega_j) \quad \because\ \text{axiom 3の(ii)} \end{aligned} \tag{A.99}$$

が得られるからである。式(2.99)の導出と同様にして、次の(ロ)の不変性も証明できる。

(ロ)(BSCのU－不変性)

TのU－不変式(2.53)が成立していれば、

$$\forall \varphi \in \Phi,\ \forall j \in J,\ BSC(U\varphi,\ \omega_j) = BSC(\varphi,\ \omega_j). \tag{A.100}$$

付録B. 類似度関数 SM を用いた任意の2つのパターン間の距離 dis と、類似度関数 SM の再帰的構成、2つの認識システム RECOGNITRON 間の認識能力の差を与える物差し DIS(1, 2；Ψ)

本付録Bでは、axiom 3を満たす式(2.22)の類似度関数 SM を用いて、3角不等式を満たすとは限らない2つの任意のパターン $\varphi, \eta \in \Phi$ 間の距離 $dis(\varphi, \eta)$ が定義できることを先ず示し、その後、dis を用いて SM´から今1つの SM が構成されることを示し、最後に、2つの認識システム RECOGNITRON 間の認識能力の差を与える物差し D(1, 2；Φ)が提案される。

B1. SM を用いた2つの任意のパターン φ, $\eta \in \Phi$ 間の距離 dis

2つの連続確率分布 p(x), q(x)間のBhattacharyya距離 B(p, q)は、

$$B(p, q) \equiv \log_e \int dx\,[p(x) \cdot q(x)]^{1/2} \tag{B.1}$$

と定義される [A19]。この事実を勘案して、Axiom 3を満たす式(2.22)の類似度関数SMを用いて、写像(距離関数)

$$dis : \Phi \times \Phi \to R^+\ (\text{非負実数全体の集合}) \tag{B.2}$$

を

$$\begin{aligned} &dis(\varphi,\ \eta) \\ &\equiv -\log_e \sum_{j \in J} [SM(\varphi,\ \omega_j) \cdot SM(\eta,\ \omega_j)]^{1/2} \end{aligned} \tag{B.3}$$

と定義すると、2つの任意のパターン $\varphi, \eta \in \Phi$ 間の距離 $dis(\varphi, \eta)$ が得られることが次の定理B.1からわかる。

[定理B.1] (**類似度関数 SM による2パターン間距離定理**)

(i)(同パターン間の0距離性) $\forall \varphi \in \Phi,\ dis(\varphi,\ \varphi) = 0.$

(ii)(対称性) $\forall \varphi \in \Phi,\ \forall \eta \in \Phi,\ dis(\varphi,\ \eta) = dis(\eta,\ \varphi).$

(iii)(非負性) $\forall \varphi \in \Phi,\ \forall \eta \in \Phi,\ dis(\varphi,\ \eta) \geqq 0.$

(iv)(非交差性パターン間の無限大距離性) $SM(\varphi,\ \omega_j) > 0$ であるようなカテゴリ番号 $j \in J$ の集合

$$J^+(\varphi\ ;\ SM) \equiv \{j \in J \mid SM(\varphi,\ \omega_j) > 0\} \tag{B.4}$$

を定義すれば、

$J^{+}(\varphi\,;\,SM) \cap J^{+}(\eta\,;\,SM) = \phi$ (the empty set) (非交差性) (B.5)

$\Rightarrow \mathrm{dis}(\varphi,\ \eta) = \infty.$ (B.6)

(v)(代表パターンとの距離) $\forall \varphi \in \Phi,\ \forall j \in J,$

$\mathrm{dis}(\omega_j,\ \eta) = -2^{-1} \cdot \log_e SM(\eta,\ \omega_j)$ (B.7)

が成立ち、よって、

$$\mathrm{dis}(\omega_j,\ \omega_k) = \begin{cases} \infty & \text{if } j \neq k \\ 0 & \text{if } j = k. \end{cases} \quad \text{(B.8)}$$

(vi)3角不等式

$\mathrm{dis}(\varphi,\ \eta) + \mathrm{dis}(\eta,\ \psi) \leqq \mathrm{dis}(\varphi,\ \psi)$ (B.9)

は必ずしも、成立しない。以下の式(B.20)が成立していれば、3角不等式(B.9)が成立する。実は、式(B.20)が成立するときに限って、3角不等式(B.9)で等号が成立し、このとき、式(B.21)が成立する。

(証明)(i)の証明：$\mathrm{dis}(\varphi,\ \varphi) =$

$$\begin{aligned} &= -\log_e \sum_{j \in J} [SM(\varphi,\ \omega_j) \cdot SM(\varphi,\ \omega_j)]^{1/2} \\ &\qquad \because \ 式(B.3) \\ &= -\log_e \sum_{j \in J} SM(\varphi,\ \omega_j) \\ &= -\log_e 1 \qquad \because \ \text{axiom 2, (ii)} \\ &= 0. \end{aligned} \quad \text{(B.10)}$$

(ii)の証明：$\mathrm{dis}(\varphi,\ \eta)$の定義式(B.3)から明らか。

(iii)の証明：

$$\begin{aligned} 0 &\leqq \sum_{j \in J} [SM(\varphi,\ \omega_j) \cdot SM(\eta,\ \omega_j)]^{1/2} \\ &= \sum_{j \in J} SM(\varphi,\ \omega_j)^{1/2} \cdot SM(\eta,\ \omega_j)^{1/2} \\ &\leqq [\sum_{j \in J} SM(\varphi,\ \omega_j)]^{1/2} \cdot \\ &\qquad [\sum_{j \in J} SM(\eta,\ \omega_j)]^{1/2} \\ &\quad \because \ \text{Schwarzの不等式} \\ &= 1^{1/2} \cdot 1^{1/2} \quad \because \ \text{axiom 2, (ii)} \\ &= 1 \end{aligned} \quad \text{(B.11)}$$

を得、よって、

$$\begin{aligned} &\mathrm{dis}(\varphi,\ \eta) \\ &= -\log_e \sum_{j \in J} [SM(\varphi,\ \omega_j) \cdot SM(\eta,\ \omega_j)]^{1/2} \\ &\geqq -\log_e 1 = 0. \end{aligned}$$

(iv)の証明：

$$\begin{aligned} &\mathrm{dis}(\varphi,\ \eta) \\ &= -\log_e \sum_{j \in J^{+}(\varphi\,;\,SM) \cap J^{+}(\eta\,;\,SM)} [SM(\varphi,\ \omega_j) \cdot \\ &\quad SM(\eta,\ \omega_j)]^{1/2} \quad \because \ 式(B.3) \\ &= -\log_e 0 \quad \because \ 式(B.5) \\ &= \infty. \end{aligned} \quad \text{(B.12)}$$

(v)の証明：

$\mathrm{dis}(\omega_k, \eta)$

$= -\log_e \sum_{j \in J} [\mathrm{SM}(\omega_k, \omega_j) \cdot \mathrm{SM}(\eta, \omega_j)]^{1/2}$

$\because$ 式(B.3)

$= -\log_e \mathrm{SM}(\eta, \omega_k)^{1/2}$ $\because$ axiom 2, (ⅰ)

を得、式(B.7)の成立がわかった。式(B.8)は式(B.7)において、$\eta = \omega_k$とおけば、axiom 2, (ⅰ)から明らかである。

(vi)の証明：

$\mathrm{dis}(\varphi, \eta) + \mathrm{dis}(\eta, \psi)$

$= -\log_e [\sum_{j \in J} \mathrm{SM}(\varphi, \omega_j)^{1/2} \cdot \mathrm{SM}(\eta, \omega_j)^{1/2}$

$\cdot \sum_{k \in J} \mathrm{SM}(\eta, \omega_k)^{1/2} \cdot \mathrm{SM}(\psi, \omega_k)^{1/2}]$

$\because$ 式(B.3)

$\leqq -\log_e \sum_{j \in J} \mathrm{SM}(\varphi, \omega_j)^{1/2} \cdot \mathrm{SM}(\psi, \omega_j)^{1/2} \cdot$

$\mathrm{SM}(\eta, \omega_j)$

$\because$ $-\log_e x$ は単調減少関数 (B.13)

$= -\log_e \sum_{j \in J^+(\varphi; \mathrm{SM}) \cap J^+(\psi; \mathrm{SM})} \mathrm{SM}(\varphi, \omega_j)^{1/2} \cdot$

$\mathrm{SM}(\psi, \omega_j)^{1/2} \cdot \mathrm{SM}(\eta, \omega_j)$ (B.14)

である。ここで

式(B.13)で等号が成立するのは、

$\forall j \in J, \forall k \in J - \{k\},$

$\mathrm{SM}(\varphi, \omega_j) \cdot \mathrm{SM}(\psi, \omega_k) \cdot \mathrm{SM}(\eta, \omega_j) \cdot \mathrm{SM}(\eta, \omega_k) = 0$ (B.15)

の場合に限ることに注意しておく。

もし、

$\exists \varepsilon \geqq 0, \forall j \in J^+(\varphi; \mathrm{SM}) \cap J^+(\psi; \mathrm{SM}),$

$1 \geqq \mathrm{SM}(\eta, \omega_j) \geqq \varepsilon$ (B.16)

であれば、不等式

$\sum_{j \in J^+(\varphi; \mathrm{SM}) \cap J^+(\psi; \mathrm{SM})} \mathrm{SM}(\varphi, \omega_j)^{1/2} \cdot$

$\mathrm{SM}(\psi, \omega_j)^{1/2} \cdot \mathrm{SM}(\eta, \omega_j)$

$\geqq \varepsilon \cdot \sum_{j \in J^+(\varphi; \mathrm{SM}) \cap J^+(\psi; \mathrm{SM})} \mathrm{SM}(\varphi, \omega_j)^{1/2} \cdot$

$\mathrm{SM}(\psi, \omega_j)^{1/2}$ (B.17)

を得るから、式(B.14)に式(B.17)を適用すれば、不等式

$\mathrm{dis}(\varphi, \eta) + \mathrm{dis}(\eta, \psi)$

$\leqq -\log_e \varepsilon + \mathrm{dis}(\varphi, \psi)$ $\because$ 式(B.12) (B.18)

が成立する。よって、3角不等式(B.9)は一般には成立しないことがわかった。

よって、式(B.18)で

$\varepsilon = 1$ (B.19)

であれば、3角不等式(B.9)が成立するが、これは、

$\exists j \in J, J^+(\varphi; \mathrm{SM}) \cap J^+(\psi; \mathrm{SM}) = \{j\}$

$\wedge \mathrm{SM}(\eta, \omega_j) = 1$ $\because$ 式(B.16) (B.20)

の場合に限ることが、axiom 2, (ⅰ)から従う。

2式(B.15), (B.20)の意味するところを勘案すれば、実は、式(B.20)が成立するときに限って、3角不等式(B.9)で等号が成立することがわかる。実際、

式(B.20)の成立

$$\Leftrightarrow$$

$$\mathrm{dis}(\varphi, \eta) = -\log_e \mathrm{SM}(\varphi, \omega_j)$$

$$\wedge \mathrm{dis}(\eta, \psi) = -\log_e \mathrm{SM}(\psi, \omega_j)$$

$$\wedge \mathrm{dis}(\varphi, \psi) = -\log_e [\mathrm{SM}(\varphi, \omega_j) \cdot \mathrm{SM}(\psi, \omega_j)] \tag{B.21}$$

が従う。 □

B2. 類似度関数 SM の再帰的構成

次の定理B.2は、式(B.3)の距離関数 dis を用いて、axiom 2を満たす式(2.22)の類似度関数 SM から axiom 2 を満たす今1つの式(B.24)の類似度関数 SM が再帰的に得られることを指摘したものである。

[定理B.2] **(パターン間距離定理による類似度関数の再帰的構成定理)**

式(B.7)の距離

$$\mathrm{dis}(\varphi, \omega_j)$$

$$= -2^{-1} \cdot \log_e \mathrm{SM}(\varphi, \omega_j)$$

$$= \log_e 1/\sqrt{\mathrm{SM}(\varphi, \omega_j)},\ j \in \mathrm{J},\ \varphi \in \Phi \tag{B.22}$$

を用いて、

$$\mathrm{SM}'(\varphi, \omega_j)$$

$$= \mathrm{dis}(\varphi, \omega_j)^{-1} / \sum_{k \in \mathrm{J}} \mathrm{dis}(\varphi, \omega_k)^{-1} \tag{B.23}$$

と定義されるの関数

$$\mathrm{SM}' : \Phi \times \Omega \to \{s \mid 0 \leqq s \leqq 1\} \tag{B.24}$$

はaxiom 2を満たす。

(証明) 先ず、

$$\sum_{k \in \mathrm{J}} \mathrm{dis}(\varphi, \omega_k)^{-1} = 0$$

$$\Leftrightarrow \forall k \in \mathrm{J},\ \mathrm{dis}(\varphi, \omega_k) = \infty$$

$$\Leftrightarrow \forall k \in \mathrm{J},\ \mathrm{SM}(\varphi, \omega_k) = 0 \quad \because \ \text{式(B.22)} \tag{B.25}$$

を得、式(B.25)はaxiom 2, (ii)に矛盾する。よって、

$$\sum_{k \in \mathrm{J}} \mathrm{dis}(\varphi, \omega_k)^{-1} > 0 \tag{B.26}$$

と仮定でき、SM´の定義式(B.23)は意味を持つ。

axiom 2, (i)の成立: 式(B.22)において、$\varphi = \omega_i$とおくと、式(2.22)の類似度関数 SM が axiom 2, (i)を満たすことから、明らかに、

$$\mathrm{dis}(\omega_i, \omega_j) = \begin{cases} 0 & \text{if } i = j \\ \infty & \text{if } i \neq j \end{cases} \tag{B.27}$$

を得る。よって、SM´の定義式(B.23)を、

$$\mathrm{SM}'(\varphi, \omega_j)$$

$$= 1/[1 + \sum_{i \in \mathrm{J} - \{j\}} \mathrm{dis}(\varphi, \omega_k)^{-1} / \mathrm{dis}(\varphi, \omega_j)^{-1}] \tag{B.28}$$

と変形した後、$\varphi = \omega_j$とおき、式(B.27)を適用すれば、

$$SM'(\omega_j, \omega_j) = 1/[1+\sum_{i \in J-\{j\}} 0/\infty] = 1 \tag{B.29}$$

を得る。また、SM´の定義式(B.23)を、

$$SM'(\varphi, \omega_j) = \mathrm{dis}(\varphi, \omega_j)^{-1} / [\mathrm{dis}(\varphi, \omega_i)^{-1} + \sum_{k \in J-\{i\}} \mathrm{dis}(\varphi, \omega_k)^{-1}] \tag{B.30}$$

と変形した後、$\varphi = \omega_i\,(i \neq j)$とおき、式(B.27)を適用すれば、

$$SM'(\omega_i, \omega_j) = 0/[\infty + \sum_{k \in J-\{i\}} 0] = 0 \tag{B.31}$$

が得られる。2式(B.29), (B.31)の成立は式(B.23)の SM´がaxiom 2,（i)を満たすことを示している。

axiom 2,（ii)の成立：SM´の定義式(B.23)から明らかである。

axiom 2,（iii)の成立：式(2.22)の類似度関数SMがaxiom 2,（iii)を満たすことから、明らかに、式(B.24)の類似度関数 SM´がaxiom 2,（iii)を満たす。 □

因みに、axiom 2を満たす式(2.22)の類似度関数 SM を以下の2定理B.3, B.4で構成しておこう。

規格化内積 nip(φ, η)とは、

$$\mathrm{nip}(\varphi, \eta) = \begin{cases} (\varphi \|\varphi\|^{-1}, \eta \|\eta\|^{-1}) & \text{if } \|\varphi\| \cdot \|\eta\| > 0 \\ 0 & \text{if } \|\varphi\| \cdot \|\eta\| = 0 \end{cases} \tag{B.32}$$

と定義され、Schwarzの不等式

$$\forall \varphi \in \mathfrak{H}, \forall \eta \in \mathfrak{H}, |\mathrm{nip}(\varphi, \eta)| \leqq 1 \tag{B.33}$$

と、

$$\mathrm{nip}(\varphi, \eta) = 1 \Leftrightarrow \exists a(\neq 0) \in Z(\text{複素数体}),\ \varphi = a \cdot \eta \neq 0 \tag{B.34}$$

との成立に注意しておく。ここで、非負規格化内積 nip$^+(\varphi, \eta)$とは、

$$\mathrm{nip}^+(\varphi, \eta) = \begin{cases} \mathrm{nip}(\varphi, \eta) & \text{if } \mathrm{nip}(\varphi, \eta) \geqq 0 \\ 0 & \text{if } \mathrm{nip}(\varphi, \eta) < 0 \end{cases} \tag{B.35}$$

と定義される。

次の定理B.3は、非負量 $-\log_e \mathrm{nip}^+(T\varphi, T\omega_j)$の逆数を規格化すれば、axiom 2を満たす式(2.22)の類似度関数 SM が得られることを指摘している。

[定理B.3]（**対数関数を利用した非負規格化内積による類似度関数の構成定理**）

$$\forall j \in J, (T\varphi, T\omega_j)\text{は実数値である} \tag{B.36}$$

の場合を考えよう。分離条件

$$\forall j \in J, \forall i \in J-\{j\}, \|T\omega_i - T\omega_j\| > 0 \tag{B.37}$$

の下で、

$$SM(\varphi, \omega_j) = \begin{cases} [-\log_e \mathrm{nip}^+(T\varphi, T\omega_j)]^{-1} \\ / \sum_{k \in J} [-\log_e \mathrm{nip}^+(T\varphi, T\omega_k)]^{-1} \end{cases}$$

$$
\begin{cases}
\quad \cdots \sum_{k \in J} [-\log_e \mathrm{nip}^+(T\varphi, T\omega_k)]^{-1} > 0 \text{ のとき} \\
p(\mathfrak{C}_j) \\
\quad \cdots \sum_{k \in J} [-\log_e \mathrm{nip}^+(T\varphi, T\omega_k)]^{-1} = 0 \text{ のとき}
\end{cases} \tag{B.38}
$$

と定義される式(2.22)の関数 SM はaxiom 2を満たす。

（証明） 3式(B.32)～(B.34)を使えば、$\mathrm{nip}^+(\varphi, \eta)$の定義式(B.35)を考慮して、

$$
-\log_e \mathrm{nip}^+(T\varphi, T\omega_j) = 0 \quad \text{if} \quad T\varphi = T\omega_j \tag{B.39}
$$

であるから、axiom 2,（i）の成立を定理B.2の証明と同様にして、示せる。axiom 2,（ii),（iii）の成立も容易に示せる。 □

次の定理B.4は、非負量 $-\log_e|\mathrm{nip}(T\varphi, T\omega_j)|^2$ の逆数を規格化すれば、axiom 2を満たす式(2.22)の類似度関数 SM が得られることを指摘している。

［定理B.4］（**対数関数を利用した規格化内積の絶対値の自乗による類似度関数の構成定理**）

分離条件式(B.37)の下で、

$$
\mathrm{SM}(\varphi, \omega_j) =
\begin{cases}
[-\log_e |\mathrm{nip}(T\varphi, T\omega_j)|^2]^{-1} \\
/ \sum_{k \in J} [-\log_e \mathrm{nip}|(T\varphi, T\omega_k)|^2]^{-1} \\
\cdots \sum_{k \in J} [-\log_e |\mathrm{nip}(T\varphi, T\omega_k)|^2]^{-1} > 0 \text{ のとき} \\
p(\mathfrak{C}_j) \\
\cdots \sum_{k \in J} [-\log_e \mathrm{nip}|(T\varphi, T\omega_k)|^2]^{-1} = 0 \text{ のとき}
\end{cases} \tag{B.40}
$$

と定義される式(2.22)の関数 SM はaxiom 2を満たす。

（証明） 2式(B.33)～(B.34)を使えば、$\mathrm{nip}(\varphi, \eta)$の定義式(B.32)を考慮して、

$$
-\log_e |\mathrm{nip}(T\varphi, T\omega_j)|^2 = 0 \quad \text{if} \quad T\varphi = T\omega_j \tag{B.41}
$$

であるから、axiom 2,（i）の成立を定理B.2の証明と同様にして、示せる。axiom 2,（ii),（iii）の成立も容易に示せる。 □

B3. 2つの認識システム RECOGNITRON 間の認識能力間の距離 DIS(1, 2 ; Φ)

パターン認識システムが処理の対象とする問題のパターン集合 Φ の変形をどう吸収するか？その吸収法には大きく分類して、

①多数ある認識システムからの選択

…あらかじめ、Φの変形の種類の各々に応じ、多数の認識システムを用意しておき、実際の変形に最も近いと思われる変形に強い RECOGNITRON を選択して、認識させる方法

②1つの認識システムの適応化

…Φ内の少数のパターン(訓練パターン)を用いてRECOGNITRON 内の構造要素を変化させ、RECOGINITRON をΦの実際の変形に適応させてから、RECOGNITRONに認識させる方法

の2通の方法がある。

これまで、T, SM, BSC 内の構造を学習で決定することにより、②の手法で、各カテゴリの代表パターンからのパターン変形を吸収研究してきたが、本章では、認識システム RECOGNITRON が多数構成されたとき、Φに最も適切な RECOGNITRON を選定することを考えよう。

この種の選定には、2つの認識システム

RECOGNITRON(1), RECOGNITRON(2) (B.42)

の、同一のパターン集合 $\Psi(\subseteq\Phi)$ に関する認識能力間の違い(距離；distance) DIS(1, 2；Ψ)を計量する方法が必要である。

n(≧1)個のパターン集合$\Psi(\subset\Phi)$を考え、Ψの元の内、RECOGNITRON(1)，RECOGNITRON(2)によって正しく認識されたパターンの集合を各々、

$$\Psi(1),\ \Psi(2)\ (\subset\Phi) \tag{B.43}$$

と表す。更に、

$\varphi\in\Psi(1)$の内、不等式

$$\max_{j\in CSF_2(\varphi, J)} SM_2(\varphi,\ \omega_j) \leq \max_{j\in CSF_1(\varphi, J)} SM_1(\varphi,\ \omega_j) \tag{B.44}$$

を満たし、かつ、RECOGNITRON(2)によって正しく認識されたパターン φ の集合を

$$\Psi(1, 2)\ (\subseteq\Psi(1)) \tag{B.45}$$

とする。同様に、$\varphi\in\Psi(2)$の内、不等式

$$\max_{j\in CSF_1(\varphi, J)} SM_1(\varphi,\ \omega_j) \leq \max_{j\in CSF_2(\varphi, J)} SM_2(\varphi,\ \omega_j) \tag{B.46}$$

を満たし、かつ、RECOGNITRON(1)によって正しく認識されたパターン φ の集合を

$$\Psi(2, 1)\ (\subseteq\Psi(2)) \tag{B.47}$$

とする。ここに、SM_j，CSF_jは各々、RECOGNITRON(j)内の、axiom 2を満たす類似度関数、カテゴリ選択関数である。このとき、選ばれているパターン集合 Ψ についてのRECOGNITRON(1)，RECOGNITRON(2)の認識能力間の距離 DIS(1, 2；Ψ)が、

$$\begin{aligned}&\mathrm{DIS}(1, 2;\Psi)\\ &\equiv 1-2^{-1}\cdot\Big[\big[\textstyle\sum_{\varphi\in\Psi(1,2)} \max_{j\in CSF_2(\varphi, J)} SM_2(\varphi,\ \omega_j)\big]/\\ &\big[\textstyle\sum_{\varphi\in\Psi(1)} \max_{j\in CSF_1(\varphi; J)} SM_1(\varphi,\ \omega_j)\big]\\ &+\big[\textstyle\sum_{\varphi\in\Psi(2,1)} \max_{j\in CSF_1(\varphi, J)} SM_1(\varphi,\ \omega_j)\big]/\\ &\big[\textstyle\sum_{\varphi\in\Psi(2)} \max_{j\in CSF_2(\varphi, J)} SM_2(\varphi,\ \omega_j)\big]\Big]\end{aligned} \tag{B.48}$$

と定義される。ここに、2^J をカテゴリ番号集合 J のすべての部分集合から成る集合として、関数

$$CSF:\ \Phi\times 2^J\rightarrow 2^J \tag{B.49}$$

は次のように定義されるカテゴリ選択関数(category-selection function)であり［B3］，［B4］、CSF_jは SM_j と、axiom 3を満たす大分類関数 BSC_j とで定義されるRECOGNITRON(j)でのカテゴリ選択関数である：

(i) $\varphi=0\vee\gamma=\phi$ の場合

$$CSF(\varphi,\ \gamma)=\phi. \tag{B.50}$$

(ii) $\varphi\neq 0\wedge\gamma\neq\phi$ の場合

$$CSF(\varphi,\ \gamma)=$$

$$\begin{cases}\{k\in\gamma\ |\ SM(\varphi,\ \omega_k)>0\}\\ \qquad \text{if } \sum_{k\in\gamma} BSC(\varphi, k)=0 & \text{(B.51)}\\ \{k\in\gamma\ |\ SM(\varphi,\ \omega_k)>0\wedge BSC(\varphi, k)=1\}\\ \qquad \text{if } \sum_{k\in\gamma} BSC(\varphi, k)>0. & \text{(B.52)}\end{cases}$$

□

上述の DIS(1, 2；Ψ)の定義式(B.44)は、音声認識システムが話者による発声の差を吸収する技術を確保するために必要とされる“話者 a と、話者 b との距離 $D_c(a, b)$”の定義式［A19］にヒントを得たものである。

次の定理B.5は、認識能力間の距離 DIS(1, 2 ; Ψ)の簡単な3性質を明らかにしたものであり、例えば、選ばれているパターン集合 Ψ についてのRECOGNITRON(1), RECOGNITRON(2)の認識能力間に差がなければ、距離 DIS(1, 2 ; Ψ)は

$$\mathrm{DIS}(1, 2\,;\,\Psi)=0 \tag{B.53}$$

であらねばならない事態が(i)によって表現されている。

[定理B.5] **(パターン集合 Ψ についてのRECOGNITRON(1), RECOGNITRON(2)の認識能力間の距離 DIS(1, 2 ; Ψ) 定理)**

(i)不等式

$$0\leqq \mathrm{DIS}(1, 2\,;\,\Psi)\leqq 1 \tag{B.54}$$

が成り立つ。

(ii)

$$[\forall\varphi\in\Psi(1),\ \max_{j\in CSF_1(\varphi, J)} SM_1(\varphi,\ \omega_j)= \max_{j\in CSF_2(\varphi, J)} SM_2(\varphi,\ \omega_j)]\wedge\Psi(1)=\Psi(1, 2) \tag{B.55}$$

かつ、

$$[\forall\varphi\in\Psi(2),\ \max_{j\in CSF_2(\varphi, J)} SM_2(\varphi,\ \omega_j) =\max_{j\in CSF_1(\varphi, J)} SM_1(\varphi,\ \omega_j)]\wedge\Psi(2)=\Psi(2, 1) \tag{B.56}$$

のとき、式(B.53)が成り立つ。

(iii) $\Psi(1, 2)=\Psi(2, 1)=\phi$ (the empty set)のとき、

$$\mathrm{DIS}(1, 2\,;\,\Psi)=1. \tag{B.57}$$

(証明) 明らかである。 □

付録C. ヒルベルト空間 $\mathfrak{H}=L_2(M,\ dm)$ でのK−L直交系

どの1つのカテゴリに帰属するどうかが判明していないパターン φ の集まりが1つのグループを形成することがある。このようなグループをクラスタ(cluster)という。処理の対象とする問題のパターン集合を幾つかのクラスタに分け、それぞれのクラスタにカテゴリ名を付与することをクラスタリング(clustering)という。通常、特徴抽出した後クラスタリングを行うことがよく行われるが、本付録Cでは、特徴抽出するときによく用いられるK−L直交系 $\{\psi_k\}_{k\in L}$ を導く。

C1.1個のサンプルパターン φ についての、圧縮近似誤差の自乗ノルム

文献

長尾真:"画像認識論", コロナ社, Feb.1983

で論じられているユークリッド空間(有限次元の数列空間) R^n での、K−L直交系 $\{\psi_k\}_{k\in L}$ の導出法を可分なヒルベルト空間(無限次元の関数空間) $\mathfrak{H}=L_2(M, dm)$ での導出法に直すことを考えよう。

1次独立な系 $\{\psi_k\}_{k\in L}$ は可分なヒルベルト空間 $\mathfrak{H}=L_2(M, dm)$ での**完全正規直交系**としよう。

このとき、パターン $\varphi\in\mathfrak{H}$ は

$$\varphi=\sum_{k\in L}(\varphi,\ \psi_k)\cdot\psi_k \tag{C.1}$$

と直交展開(フーリェ式展開)されるが、$L'\subseteq L$ を満たす集合 L の部分集合 L′を選び、φ から残余

$$\varphi_\perp\equiv\sum_{k\in L-L'}(\varphi,\ \psi_k)\cdot\psi_k \tag{C.2}$$

を差し引いたパターン

$$\varphi' = \varphi - \varphi_{\perp}$$
$$= \sum_{k \in L'} (\varphi,\ \psi_k) \cdot \psi_k$$
ここに、$(\varphi',\ \varphi_{\perp}) = 0$ (C.3)

を φ の代りと考えよう。高次元パターン φ を低次元パターン φ' で近似することになり、次元 $|L|$ を次元 $|L'|$ に圧縮することを考えていることになる。

ここで、

$$h(x, y) = \sum_{k \in L - L'} \psi_k(x) \cdot \overline{\psi_k}(y) \tag{C.4}$$

を核関数とし、

$$(H\eta)(x) = \int_M dm(y)\ h(x, y) \cdot \eta(y)$$
$$= \sum_{k \in L - L'} \psi_k(x) \cdot (\eta,\ \psi_k)$$
$$\because\ 2式 (C.4),\ (2.5)$$
for any $x \in M$ (C.5)

と定義されるする積分作用素 H を定義すれば、次の補助定理C.1が成り立ち、**Hは正値作用素であること**がわかる。

[補助定理C.1]

$$\forall \varphi \in \mathfrak{H},\ 0 \leqq \|\varphi_{\perp}\|^2 = (H\varphi,\ \varphi). \tag{C.6}$$

(証明) $\|\varphi_{\perp}\|^2$

$$= (\varphi_{\perp},\ \varphi_{\perp})$$
$$= \sum_{k \in L - L'} (\varphi,\ \psi_k) \cdot \overline{(\varphi,\ \psi_k)} \quad \because\ 式 (C.2)$$
$$= \sum_{k \in L - L'} \left[\int_M dm(x)\ \varphi(x) \cdot \overline{\psi_k}(x)\right] \cdot$$
$$\left[\int_M dm(y)\ \overline{\varphi}(y) \cdot \psi_k(y)\right] \quad \because\ 式 (2.5)$$
$$= \sum_{k \in L - L'} \int_M dm(y) \int_M dm(x)$$
$$\varphi(x) \cdot \overline{\psi_k}(x) \cdot \overline{\varphi}(y) \cdot \psi_k(y) \tag{C.7}$$
$$= \int_M dm(y)\ \overline{\varphi}(y) \cdot \int_M dm(x)$$
$$\left[\sum_{k \in L - L'} \psi_k(y) \cdot \overline{\psi_k}(x)\right] \cdot \varphi(x)$$
$$= \int_M dm(y)\ \overline{\varphi}(y) \cdot \int_M dm(x)\ h(y, x) \cdot \varphi(x) \quad \because\ 式 (C.4)$$
$$= (H\varphi,\ \varphi). \quad \because\ 式 (C.5) \qquad \square$$

一方、

$$q_{\varphi}(x, y) \equiv \varphi(x) \cdot \overline{\varphi}(y) \tag{C.8}$$

を核関数とし、

$$(Q_{\varphi}\eta)(x)$$
$$\equiv \int_M dm(y)\ q_{\varphi}(x, y) \cdot \eta(y)$$
$$= \varphi(x) \cdot (\eta,\ \varphi) \quad \because\ 2式 (C.8),\ (2.5)$$
for any $x \in M$ (C.9)

と定義されるする積分作用素 Q_{φ} を考えよう。このとき、次の補助定理C.2が成立し、φ を φ' に圧縮し、φ を φ' で近似するときの誤差 $\varphi_{\perp}$ の自乗ノルム $\|\varphi_{\perp}\|^2$ が、補助定理C.1と異なる形式で再表現されることになる。

[補助定理C.2]

$$\forall \varphi \in \mathfrak{H},\ 0 \leqq \|\varphi_{\perp}\|^2 = \sum_{k \in L - L'} (Q_{\varphi}\psi_k,\ \psi_k). \tag{C.10}$$

（証明） 式(C.7)から出発し、

$\| \varphi_{\perp} \|^2$

$= \sum_{k \in L-L'} \int_M dm(y) \int_M dm(x)$

$\varphi(x) \cdot \overline{\psi_k}(x) \cdot \overline{\varphi}(y) \cdot \psi_k(y)$ ∵ 式(C.7)

$= \sum_{k \in L-L'} \int_M dm(x)\ \overline{\psi_k}(x) \cdot \int_M dm(y)$

$q_\varphi(x, y) \cdot \psi_k(y)$ ∵ 式(C.8)

$= \sum_{k \in L-L'} \int_M dm(x)\ \overline{\psi_k}(x) \cdot (Q_\varphi \psi_k)(x)$

∵ 式(C.9)

$= \sum_{k \in L-L'} (Q_\varphi \psi_k, \psi_k).$ ∵ 式(2.5) □

さて、

$\forall \eta \in \mathfrak{H},$

$(Q_\varphi \eta, \eta)$

$= (\eta, \varphi) \cdot (\varphi, \eta)$ ∵ 式(C.9)

$|(\eta, \varphi)|^2 \geqq 0$ (C.11)

を得、式(C.9)の線形作用素 Q_φ も正値作用素であることがわかる。

C2. 多数のサンプルパターン φ_n についての、圧縮近似誤差の自乗ノルム

確率条件

$[\forall n (=1, 2, \cdots, n),\ 0 < p_n < 1]$ (C.12)

$\wedge \sum_{n=1}^{N} p_n = 1$ (C.13)

を満たす生起確率 $p_n (n=1, 2, \cdots, N)$ についての、量 … の期待値

$E(\cdots) \equiv \sum_{n=1}^{N} p_n \cdots$ (C.14)

を導入しよう。

N 個のサンプルパターン $\varphi_n (n=1, 2, \cdots, N)$ について、C1章の近似式(C.3)を行うとしよう：

$\varphi_n = \sum_{k \in L} (\varphi_n, \psi_k) \cdot \psi_k$ (C.15)

と直交展開されるが、$L' \subseteq L$ を満たす集合 L の部分集合 L′を選び、φ_n から残余

$\varphi_{n\perp} \equiv \sum_{k \in L-L'} (\varphi_n, \psi_k) \cdot \psi_k$ (C.16)

を差し引いたパターン

$\varphi_n' = \varphi_n - \varphi_{n\perp}$

$= \sum_{k \in L'} (\varphi_n, \psi_k) \cdot \psi_k.$ (C.17)

を考えよう。

このとき、式(C.8)の $q_\varphi(x, y)$ は、

$q(x, y)$

$= E[\varphi_n(x) \cdot \overline{\varphi_n}(y)\ ;\ n=1, 2, \cdots, N]$

$= E[q_n(x, y)\ ;\ n=1, 2, \cdots, N]$

$= \sum_{n=1}^{N} p_n \cdot q_n(x, y)$

$= \sum_{n=1}^{N} p_n \cdot \varphi_n(x) \cdot \overline{\varphi_n}(y)$ (C.18)

と変り、この $q(x, y)$ を核関数とし、

$$\begin{aligned}&(Q\eta)(x)\\&=\int_M dm(y)\ q(x, y)\cdot\eta(y)\\&=\sum_{n=1}^{N} p_n\cdot\varphi_n(x)\cdot(\eta, \varphi_n)\\&\qquad \because\ 2式(C.18),\ (2.5)\\&\quad \text{for any } x\in M\end{aligned}\tag{C.19}$$

と定義されるする積分作用素 Q を考えよう。

各サンプルパターン φ_n を $\varphi_n'(=\varphi_n-\varphi_{n\perp})$ に圧縮し、各 φ_n を φ_n' で近似したときの誤差 $\varphi_{n\perp}$ の自乗ノルム $\|\varphi_{n\perp}\|^2$ を $n=1, 2, \cdots, N$ にわたり平均して得られる **“サンプルパターン1個当りの圧縮近似誤差の自乗ノルムの期待値”** $E[\|\varphi_{n\perp}\|^2; n=1, 2, \cdots, N]$ は、補助定理C.2より、

$$\begin{aligned}&0\leqq E\ [\|\varphi_{n\perp}\|^2; n=1, 2, \cdots, N]\\&\equiv\sum_{n=1}^{N} p_n\cdot\|\varphi_{n\perp}\|^2\\&=\sum_{n=1}^{N} p_n\cdot\sum_{k\in L-L'}(Q_{\varphi n}\psi_k,\ \psi_k)\\&=\sum_{k\in L-L'}(Q\psi_k,\ \psi_k).\end{aligned}\tag{C.20}$$

と表されることに注意しておく。同時に、作用素 Q は、正値作用素 $Q\varphi_n$ の $n=1, 2, \cdots, N$ にわたる期待値として、

$$\begin{aligned}&Q=E[Q_{\varphi n}; n=1, 2, \cdots, N]\\&=\sum_{n=1}^{N} p_n\cdot Q_{\varphi n}\end{aligned}\tag{C.21}$$

と表されていることにも、注意しておく。

C3. K－L直交系の導出と、その圧縮に関する最適性

先ず、

$$\begin{aligned}&\forall\eta\in\mathfrak{H},\ (Q\eta,\ \eta)\\&=\int_M dm(x)\left[\int_M dm(y)\ q(x, y)\cdot\eta(y)\right]\cdot\\&\quad\overline{\eta}(x)\quad\because\ 2式(C.19),\ (2.5)\\&=\sum_{n=1}^{N} p_n\cdot\int_M dm(x)\left[\int_M dm(y)\ \varphi_n(x)\cdot\right.\\&\left.\overline{\varphi_n}(y)\cdot\eta(y)\right]\cdot\overline{\eta}(x)\quad\because\ 式(C.18)\\&=\sum_{n=1}^{N} p_n\cdot\int_M dm(x)\ \varphi_n(x)\cdot\overline{\eta}(x)\cdot\\&\left[\int_M dm(y)\ \overline{\varphi_n}(y)\cdot\eta(y)\right]\\&=\sum_{n=1}^{N} p_n\cdot(\varphi_n,\ \eta)\cdot(\eta,\ \varphi_n)\\&=\sum_{n=1}^{N} p_n\cdot|(\eta,\ \varphi_n)|^2\end{aligned}\tag{C.22}$$

$$\geqq 0\tag{C.23}$$

が成立し、2式(C.19), (C.20)の線形作用素 Q は正値であることがわかる。因みに、式(C.22)は $(Q\eta, \eta)$ がS.Suzukiの唱えた“生起確率 pn を持つサンプルパターン φ_n の系列とパターン η との間の**平均類似度**(degree of average similarity) [B1], [B5]”であることがわかる。

正値作用素Qの固有値方程式

$$Q\varphi=\lambda\varphi\tag{C.24}$$

の解である固有値λ，固有ベクトル$\varphi(\neq 0)$につき，このようなλの集まりの下限，上限を各々，

$$\inf \lambda, \sup \lambda \tag{C.25}$$

と表す.

次の2補助定理C.3，C.4は，自己共役作用素Gのスペクトル理論からよく知られている。

非負値2次汎関数$(Q\varphi, \varphi)$のとる値(測度的ユニタリ不変量［B1]，[B5])についての，下限·上限を明らかにしているのが、次の2補助定理C.3である。

［補助定理C.3］(**非負2次汎関数$(Q\varphi, \varphi)$の下限·上限**)

正値作用素Qから定まる2次汎関数$(Q\varphi, \varphi)$について，

$$\forall \varphi \in \mathfrak{H},$$
$$0 \leqq [\inf \lambda]\cdot(\varphi, \varphi) \leqq (Q\varphi, \varphi) \leqq [\sup \lambda]\cdot(\varphi, \varphi). \tag{C.26}$$

□

次の補助定理C.4は、2式(C.19)，(C.20)の線形作用素 Q の固有値λ_ℓの列が式(C.27)の如く、整列することなどを指摘している。

［補助定理C.4］(正値作用素Qの固有値定理)

2式(C.19)，(C.20)の線形作用素 Q の，すべての固有値の集合は高々可算個の非負実数値からなる点スペクトル系であり，固有値方程式

$$Q\eta_\ell = \lambda_\ell\cdot\eta_\ell \wedge \|\eta_\ell\| = 1,\ \ell \in L \tag{C.27}$$

の解である“第$\ell \in L$番目の固有値λ_ℓと，(λ_ℓに属する)ノルム規格化固有ベクトルη_ℓ”とを考えると，その大きさの順に

$$\lambda_1 \geqq \lambda_2 \geqq \cdots \geqq 0 \tag{C.28}$$

と並べることができる。但し，重複を許して，重複の回数だけ固有値を並べ，1つの固有値に属する固有ベクトルは唯1個に限るものとする。このとき，各η_kの正規直交性

$$(\eta_k, \eta_\ell) = 1 \ \text{if}\ k = \ell,\ = 0\ \text{if}\ k \neq \ell \tag{C.29}$$

も成り立っており，正値作用素 Q がその各固有値λ_k，各固有ベクトルη_kによってスペクトル分解されるというスペクトル表現

$$\forall \varphi \in \mathfrak{H},\ Q_\varphi = \sum_{\ell \in L} \lambda_\ell\cdot(\varphi, \eta_\ell)\cdot\eta_\ell \tag{C.30}$$

が成り立つ.

□

ここで、2変数 x, y の関数

$$p(x, y) \equiv \sum_{\ell \in L} \lambda_\ell\cdot\eta_\ell(x)\cdot\overline{\eta_\ell}(y) \tag{C.31}$$

を投入すれば、式(C.18)の2変数 x, y の関数 $q(x, y)$を積分核に持つ式(C.19)の積分作用素 Q は、

$$\begin{aligned}
&\forall \eta \in \mathfrak{H},\\
&(Q\eta)(x)\\
&= \sum_{\ell \in L} \lambda_\ell\cdot(\eta, \eta_\ell)\cdot\eta_\ell(x) \quad \because \text{式(C.30)}\\
&= \sum_{\ell \in L} \lambda_\ell\cdot\left[\int dm(y)\,\eta(y)\cdot\overline{\eta_\ell}(y)\right]\cdot\eta_\ell(x) \qquad \because \text{式(2.5)}\\
&= \int dm(y)\,\eta(y)\cdot\left[\sum_{\ell \in L} \lambda_\ell\cdot\eta_\ell(x)\cdot\overline{\eta_\ell}(y)\right]\\
&= \int dm(y)\,p(x, y)\cdot\eta(y)\\
&\text{for any}\ \ x \in M
\end{aligned} \tag{C.32}$$

と再表現される。

$$\forall x \in M,\ \forall y \in M,\ q(x, y) = p(x, y) \tag{C.33}$$

であるということになる。式(C.31)の核関数 p は式(C.18)の核関数 q のスペクトル表現と称され

てよい。尚、式(C.31)の p，式(C.32)の Q は各々、式(C.4)の h，式(C.5)の H に対応する表現である。

さて、非負量($Q\varphi\|\varphi\|^{-1}, \varphi\|\varphi\|^{-1}$)を最大ならしめるのは，式(C.26)からわかるように、φ が $\varphi=\eta_1$ のときである。この事実などから、以下の主張が成り立つ：式(C.20)から、

$$E[\|\varphi_{n\perp}\|^2;\ n=1, 2, \cdots, N]$$
$$=\sum_{k\in L-L'}\lambda_k \quad \because \text{ 式(C.27)} \tag{C.34}$$

が成立しているが、式(C.28)のごとく、各固有値 λ_k は大から小の順に整列化されているから、圧縮に伴う近似誤差の自乗ノルム $E[\|\varphi_{n\perp}\|^2;\ n=1, 2, \cdots, N]$ が最小となっている。 □

完全正規直交系 $\{\psi_k\}_{k\in L}$ の部分系 $\{\psi_k\}_{k\in L}$ として、2式(C.19)，(C.20)の線形作用素 Q の固有ベクトル系 $\{\eta_k\}_{k\in L'}$ をとることができる。2式(C.19)，(C.20)の線形作用素 Q の固有ベクトル系 $\{\eta_k\}_{k\in L'}$ はK－L直交系(karhunen-Loêve orthogonal system)であるといわれる。正規直交をなす部分系 $\{\psi_k\}_{k\in L'}$ として、K－L直交系以外の正規直交系を採用すれば、式(C.20)の $E[\|\varphi_{n\perp}\|2;\ n=1, 2, \cdots, N]$ はその最小値より大きくなる。

［**結論**］

式(C.3)の直交展開 φ'の正規直交系 $\{\psi_k\}_{k\in L'}$ として、K－L直交系 $\{\eta_k\}_{k\in L'}$ をとることにより、$|L|$次元を$|L'|$次元に下げたときの、圧縮に伴う近似誤差の自乗ノルム(平均自乗誤差)$E[\|\varphi_{n\perp}\|^2;\ n=1, 2, \cdots, N]$ を最小にすることが出来る。 □

付録D. 有限個の振幅値をとるモデル構成作用素 T

本付録Dでは、2.1節の定理2.1を適用し、対【Φ, T】が得られるような**有限個の振幅値をとるパターンモデル Tφ** を説明しよう。

不等式系

$$\forall x\in M,$$
$$-1=e_{2p+1}^-(x)<e_{2p}^-(x)<\cdots<e_2^-(x)$$
$$<e_1^-(x)<0<e_1^+(x)$$
$$<e_2^+(x)<\cdots<e_{2p}^+(x)<e_{2p+1}^+(x)=+1 \tag{D.1}$$

を満たす閾値関数 $e_k^\pm(x)$の組

$$e_k^\pm(x),\ x\in M, k=1, 2, \cdots, 2p+1 \tag{D.2}$$

と，不等式系

$$\forall x\in M,$$
$$-1=t_{2p+1}^-(x)<t_{2p}^-(x)<\cdots<t_2^-(x)$$
$$<t_1^-(x)<0<t_1^+(x)$$
$$<t_2^+(x)<\cdots<t_{2p}^+(x)<t_{2p+1}^+(x)=+1$$

ここに、

$$e_{k-1}^+(x)<t_k^+(x)\leqq e_k^+(x)$$
$$e_k^-(x)\leqq t_k^-(x)<e_{k-1}^-(x) \quad (2\leqq k\leqq 2p+1) \tag{D.3}$$

を満たす関数 $t_k^\pm(x)$の組

$$t_k^\pm(x),\ x\in M, k=1, 2, \cdots, 2p+1 \tag{D.4}$$

とを用意する。$(T\varphi)(x)$を次のように実践する：

$$
(T\varphi)(x)=
\begin{cases}
t_k^{+}(x) \\
\quad \text{if } e_{k-1}^{+}(x) < \varphi(x)/\sup_{y\in M}|\varphi(y)| \\
\qquad \leqq e_k^{+}(x)\ (2\leqq k\leqq 2p+1) \\
0 \\
\quad \text{if } e_1^{-}(x) \leqq \varphi(x)/\sup_{y\in M}|\varphi(y)| \leqq e_1^{+}(x) \\
t_k^{-}(x) \quad \text{if } e_k^{-}(x) \leqq \varphi(x)/\sup_{y\in M}|\varphi(y)| \\
\qquad < e_{k-1}^{-}(x)\ (2\leqq k\leqq 2p+1)
\end{cases}
$$

$$\text{for any } x\in M \tag{D.5}$$

□

$T\varphi$の定義式(D.5)には$t_1^{\pm}(x)$は使われていないことに注意しよう。このとき次の2補助定理D.1, D.2が成り立つ。

［補助定理D.1］（**零パターンモデル定理**）

式(D.5)で定義されている式(2.1)の写像 T に対し、

$$
\begin{aligned}
&\forall x\in M, \\
&e_1^{-}(x) \leqq \varphi(x)/\sup_{y\in M}|\varphi(y)| \leqq e_1^{+}(x) \\
&\Leftrightarrow T\varphi=0.
\end{aligned} \tag{D.6}
$$

□

［補助定理D.2］（**パターンモデルの不動点定理**）

式(D.5)で定義されている式(2.1)の写像 T に対し、振幅の絶対値の上限の“0, 1”規格化条件

$$\sup_{y\in M}|\eta(y)| \in \{0, 1\} \tag{D.6}$$

の下で、

$$
\begin{aligned}
&[\forall x\in M, \exists k\in\{k\mid 2\leqq k\leqq 2p+1\}, \\
&\eta(x)\in\{t_k^{\pm}(x)],\ 0\} \\
&\Rightarrow T\eta=\eta.
\end{aligned} \tag{D.7}
$$

□

このとき，次の定理D.1が成り立ち，離散有限個の振幅値をとるパターンモデル$T\varphi$が得られた.

［定理D.1］（**離散有限個の振幅値をとるモデル構成作用素Tの構成定理**）

不等式(D.1)を満たす閾値関数$e_k\pm(x)$の，式(D.2)の組と，不等式(D.3)を満たす関数$t_k\pm(x)$の，式(D.4)の組とを用意すれば,式(D.5)のように定義された式(2.1)の写像Tは2.1節のaxiom 1の(ｉ), (ii)，(iii)の3後半を満たし，補助定理D.2から，2.1節のaxiom 1の(iv)をも満たすようにΦをとることができる. □

例えば、不等式(D.3)を満たす関数$t_k\pm(x)$の，式(D.4)の組を，簡単には,

$$
\begin{aligned}
&t_{2p+1}^{+}(x)=1,\ t_{2p+1}^{-}(x)=-t_{2p+1}^{+}(x) \\
&t_k^{+}(x)=2^{-1}\cdot[e_{k-1}^{+}(x)+e_k^{+}(x)] \\
&t_k^{-}(x)=-t_k^{+}(x) \\
&(2\leqq k\leqq 2p+1)\ \text{for any } x\in M
\end{aligned} \tag{D.8}
$$

と選ぶことができるが，p=1 と選定しているとき，式(D.5)のパターンモデル $T\varphi$ は,

$$(T\varphi)(x)=$$

$$\begin{cases} 1 & \text{if } 2/3 < \varphi(x)/\sup_{y\in M}|\varphi(y)| \leqq 1 \\ 1/2 & \text{if } 1/3 < \varphi(x)/\sup_{y\in M}|\varphi(y)| \leqq 2/3 \\ 0 & \text{if } -1/3 \leqq \varphi(x)/\sup_{y\in M}|\varphi(y)| \leqq 1/3 \\ -1/2 & \text{if } -2/3 \leqq \varphi(x)/\sup_{y\in M}|\varphi(y)| < -1/3 \\ -1 & \text{if } -1 \leqq \varphi(x)/\sup_{y\in M}|\varphi(y)| < -2/3 \end{cases}$$

$$\text{for any } x \in M \qquad \text{(D.9)}$$

である。

付録E. 座標値間の相関を考慮したaxopm 2を満たす類似度関数 SM の構成

本付録Eでは、axiom 2を満たす式(2.22)の類似度関数 SM が2つの座標値 x, y 間の相関の度合いを考慮して、構成される。

2つの座標値 x, y 間で、2つのパターンモデル $T\varphi(x)$, $T\omega_j(y)$が似ていても、x, y 間がもともと相関が低ければ、 $T\varphi(x)$, $T\omega_j(y)$の似ている程度を抑えようという機能を持つ類似度関数 SM が構成される。

文献

井上光平，浦浜喜一：“多変量写像法による任意形状クラスタの抽出”，電子情報通信学会論文誌D-Ⅱ, vol.J84－D-Ⅱ，no.2, pp.229-237, Feb.2001

での、2つの画素 i, j 間の類似度 w_{ij} の形式を勘案し、2つの座標値(画素)

$$x = \mathrm{col}(x_1\ x_2 \cdots x_q)\ (\text{列ベクトル}) \in M$$

$$y = \mathrm{col}(y_1\ y_2 \cdots y_q) \in M \qquad \text{(E.1)}$$

においてーで、2つのパターンモデル $T\varphi(x)$, $T\omega_j(y)$の似ている程度を、

$$S'(T\varphi, T\omega_j)(x) \equiv \int_M dm(y) \exp[-a_j(1)^{-1} \cdot |(T\varphi)(x) - (T\omega_j)(y)|^2 - -a(2)^{-1} \cdot \rho_j(|x-y|)] / \int_M dm(z) \qquad \text{(E.2)}$$

と、想定してみよう。ここに、

座標点 x, y 間の内積 [x, y]

$$\equiv \sum_{i=1}^{q} x_i \cdot y_i \qquad \text{(E.3)}$$

座標点 x の原点からの距離 $|x| \equiv [[x, x]]^{1/2}$ (E.4)

であって、このとき、

$$\exp[-a(2)^{-1} \cdot \rho_j(|x-y|)] \qquad \text{(E.5)}$$

は、座標点 y から眺めた座標点 x への寄与の度合いを表しており、この寄与の度合いを、2つのパターンモデル $T\varphi(x)$, $T\omega_j(y)$の似ている程度

$$\exp[-a_j(1)^{-1} \cdot |(T\varphi)(x) - (T\omega_j)(y)|^2 \qquad \text{(E.6)}$$

に加味して、座標点 x での$S'(T\varphi, T\omega_j)(x)$に取り入れていることに注意する。

$$\forall j \in J,\ a_j(1)^{-1} > a(2)^{-1} > 0 \qquad \text{(E.7)}$$

と選ぶことが望ましい。何故ならば、$(T\varphi)(x)$, $(T\omega_j)(y)$間の少しの違い$|(T\varphi)(x)-(T\omega_j)(y)|$をx, y間の少しの違い$|x-y|$より重要視するためである。例えば、

$$a_j(1)^{-1}=2000,\ a(2)^{-1}=0.05 \tag{E.8}$$

と選べばよい。

登場している非負実数値関数$\rho(|x-y|)$は、2つの座標値x, y間の相関の度合いを対称化したものであり、1実非負変数$u(\geqq 0)$の非負実数値関数$\rho_j(u)$として、

$$(\text{イ})\ \rho_j(u)=b_j\cdot u^2,\ b_j>0 \tag{E.9}$$

$$(\text{ロ})\ \rho_j(u)=b_j(1)\cdot\{1-\exp[b_j(0)^{-1}\cdot u^2]\},\quad b_j(0),\ b_j(1)>0 \tag{E.10}$$

などが考えられる。関数$\rho_j(u)$に要求されるのは、次の3性質(一), (二), (三)である：

(一) $\rho_j(0)=0.$ (E.11)

(二)ある正定数$(\infty>)c_j(>0)$が存在して、$u>c_j$のとき、$\rho_j(u)$は変数uに関し単調増加である。

(三)無限大になってもよいある正定数$(\infty\geqq)d_j(>0)$が存在して、

$$\lim_{u\to\infty}\rho_j(u)=d_j. \tag{E.12}$$

□

明らかに、式(E.2)の$S'(T\varphi, T\omega_j)(x)$について、不等式

$$\forall x\in M,\ 0\leqq S'(T\varphi, T\omega_j)(x)\leqq 1 \tag{E.13}$$

が成立する。$S'(T\varphi, T\omega_j)(x)$を積分して、

$$S(\varphi, \omega_j) \equiv \int_M dm(x)S'(T\varphi, T\omega_j)(x)/\int_M dm(y) \tag{E.14}$$

を導入する。明らかに、式(E.14)の$S(\varphi, \omega_j)$について、不等式

$$0\leqq S(\varphi, \omega_j)\leqq 1 \tag{E.15}$$

が成立する。Tの下での不変性

$$\forall\varphi\in\Phi,\ \forall j\in J,\ S(T\varphi, T\omega_j) =S(T\varphi, \omega_j)=S(\varphi, T\omega_j)=S(\varphi, \omega_j) \quad \because\ \text{axiom 1, (iii)の後半} \tag{E.16}$$

も成立している。ここで、2つの閾値$s_0(j)$, $s_1(j)$の、$j\in J$にわたる系

$$s_0(j),\ s_1(j),\ j\in J \tag{E.17}$$

を、不等式

$$\max_{i\in J-\{j\}} S(\omega_i, \omega_j)\leqq s_0(j) <s_1(j)\leqq S(\omega_j, \omega_j) \tag{E.18}$$

を満たすように式(2.19)の代表パターンω_jの集合Ωと、座標値間の相関を与える1実変数uの各負実数値関数$\rho_j(u)$ $(j\in J)$を導入できるとしよう。その後、式(E.12)の$S(\varphi, \omega_j)$を、

$$s_{01}(\varphi, \omega_j)= \begin{cases} 1 & \text{if } s_1(j)\leqq S(\varphi, \omega_j) \\ [S(\varphi, \omega_j)-s_0(j)]\ /\ [s_1(j)-s_0(j)] & \text{if } s_0(j)<S(\varphi, \omega_j)<s_1(j) \\ 0 & \text{if } S(\varphi, \omega_j)\leqq s_0(j) \end{cases} \tag{E.19}$$

へと変換する。その後、$s_{01}(\varphi, \omega_j)$を規格化して、

$SM(\varphi, \omega_j)=$

$$\begin{cases} s_{01}(\varphi,\ \omega_j)/\sum_{k\in J} s_{01}(\varphi,\ \omega_k) \\ \qquad \text{if } \sum_{k\in J} s_{01}(\varphi,\ \omega_k)>0 \\ p(\mathfrak{C}_j)\ \text{if } \sum_{k\in J} s_{01}(\varphi,\ \omega_k)=0 \end{cases} \tag{E.20}$$

を定義する。このとき、次の定理E.1が成立し、座標値間の相関を取り入れ、axiom 2を満たす式(2.22)の類似度関数 SM が構成できた。

［定理E.1］（**座標値間の相関 ρ_j を取り入れた類似度関数 SM の構成定理**）

式(E.20)の如く、定義された式(2.22)の関数 SM はaxiom 2を満たす。

（証明） axiom 2，（ｉ）の成立： 式(E.19)の $s_{01}(\varphi,\ \omega_j)$ において、$\varphi=\omega_i$ とおくと、不等式(E.18)を仮定していることから、明らかに、

$$s_{01}(\omega_i,\ \omega_j)= \begin{cases} 1 & \text{if } i=j \\ 0 & \text{if } i\neq j \end{cases} \tag{E.21}$$

を得る。よって、SM の定義式(E.20)を適用すれば、

$$\begin{aligned} &SM'(\omega_j,\ \omega_j) \\ &=s_{01}(\omega_j,\ \omega_j)/[s_{01}(\omega_j,\ \omega_j)+\sum_{k\in J-\{j\}} \\ &s_{01}(\omega_i,\ \omega_j)] \\ &=1 \end{aligned} \tag{E.22}$$

を得る。また、任意の $i\in J-\{j\}$ について、

$$\begin{aligned} &SM(\omega_i,\ \omega_j) \\ &=s_{01}(\omega_i,\ \omega_j) \\ &/[s_{01}(\omega_i,\ \omega_i)+\sum_{k\in J-\{i\}} s_{01}(\omega_i,\ \omega_k)] \\ &=0 \end{aligned} \tag{E.23}$$

が得られる。

axiom 2，（ii）の成立：SM の定義式(E.20)から明らかである。

axiom 2，（iii）の成立：T の下での不変性

$$\forall\varphi\in\Phi,\ \forall j\in J,\ s_{01}(T\varphi,\ \omega_j)=s_{01}(\varphi,\ \omega_j)$$

2式(E.16)，(E.19) (E.24)

$$\therefore\quad SM(T\varphi,\ \omega_j)=SM(\varphi,\ \omega_j) \tag{E.25}$$

が成立し、式(E.20)の類似度関数 SM がaxiom 2，（iii）を満たす。 □

付録F. パターン認識の基本的手法

分類操作の対象となりうるものをパターンという。

この種の分類操作は知能の働きの源泉である。本付録 F では、パターン認識技術の基本に立ち戻り、変形しているパターン φ の特徴量の組を変換･整形化し、分類する基本的な手法を解説する。

F1. 不動点を連想してパターン認識を行う手法

axiom 1を満たす対【Φ, T】を導入し、パターン変換作用素

$$G:\Phi\to\Phi \tag{F.1}$$

を用意して、処理の対象とする問題のパターン $\varphi \in \Phi$ のモデル $T\varphi$ を

$$TGT\varphi \in \Phi \tag{F.2}$$

へと変換したとき、得られたこのパターン $\eta \equiv GT\varphi \in \Phi$ のモデル $T\eta = TGT\varphi \in \Phi$ を φ から想起されるパターン想起作用素 G による像という。この際、モデル $T\varphi$ を見たり聞いたりしたならば、原パターン φ かのように見えたり聞こえたりすることを保証する **"同一知覚原理"** が成立しているようなaxiom 1を満たす式(2.1)のモデル構成作用素 T が選ばれている必要がある。

パターン想起作用素の列

$$G_0, G_1, \cdots, G_t \tag{F.3}$$

をその都度、適切に選定し、パターン φ を

(1#)(初期段階；initialization)　$\varphi_0 \equiv T\varphi$ (F.4)

(2#)(帰納段階；induction)

$$\varphi_{s+1} \equiv TG_s T\varphi_s,\ s=0, 1, 2, \cdots, t \tag{F.5}$$

という形式で多段階変換し、不動点方程式

$$\varphi_{t+1} = \varphi_t,\ \text{つまり、}\ TG_t T\varphi_t = \varphi_t \tag{F.6}$$

の成立を終了条件とする多段階認識過程

$$\varphi_0, \varphi_1, \cdots, \varphi_t \tag{F.7}$$

に関する研究については、2文献［B3］,［B4］にある。このとき、

φ_t は φ から連想されるパターンモデルであり、
φ の知覚的記憶表象(φ を知覚して、短期記憶内に
想起･形成された表象)である (F.8)

といわれる。

不動点方程式(F.6)が成立すれば、正常なパターン(変形程度が少なくて、唯1つの、あるカテゴリ $\mathfrak{C}_j$ に帰属しているパターン) φ に対しては、φ_t は φ の帰属する(第$j \in J$番目の)カテゴリ $\mathfrak{C}_j$ の持つ諸性質を典型的に代表しているパターン $\omega_j \in \Omega \subset \Phi$ のモデル$T\omega_j$になっているように、式(F.3)のパターン想起作用素 Gs の列が選ばれることが望ましい［B3］,［B4］。

F2. 1 次独立な系 $\{\psi_\ell\}_{\ell \in L}$ に基づく特徴抽出と、パターンモデル $T\varphi$

一般に、可分な一般抽象ヒルベルト空間 $\mathfrak{H}$ の元からなる系 $\{\psi_k\}_{k \in L}$ は、**1次独立**であるとしよう。このとき、処理の対象とするパターン $\varphi \in \Phi \subset \mathfrak{H}$ が、

第$k \in L$番目の**パターン形状素**と呼ばれる ψ_k から
なる系($\mathfrak{H}$ の基底の一部) $\psi_k, k \in L$ の線形1次
結合 $\sum_{k \in L} a_k \cdot \psi_k$ で近似される場合の近似誤差

$$\| \varphi - \sum_{k \in L} a_k \cdot \psi_k \| \tag{F.9}$$

を最小ならしめる各複素係数 $a_k \equiv a_k(\varphi)$ については、**最小自乗法**(method of least squares)を適用して得られる連立1次方程式

$$\sum_{m \in L} a_m(\varphi) \cdot (\psi_m, \psi_k) = (\varphi, \psi_k),\ k \in L \tag{F.10}$$

を解いて求めることが出来る。この時、

$$\forall k \in L,\ (\varphi_\perp, \psi_k) = 0 \tag{F.11}$$

を満たす $\varphi_\perp \in \mathfrak{H}$ が存在して、原パターン φ の表現

$$\varphi = \sum_{k \in L} a_k(\varphi) \cdot \psi_k + \varphi_\perp \tag{F.12}$$

が成り立つから、パターン φ から抽出される第 $\ell \in L$ 番目の特徴量 $u(\varphi, \ell)$として、

$$u(\varphi, \ell) \equiv \begin{cases} 0 & \cdots \forall k \in L, a_k(\varphi) = 0 \text{ のとき} \\ a_k(\varphi) / [\sup_{k \in L} | a_k(\varphi) |] & \cdots \exists k \in L, a_k(\varphi) \neq 0 \text{ のとき} \end{cases} \quad (F.13)$$

を採用してみよう。

例えば、内積

$$(\varphi, \eta) = \Sigma_{x_1=0, \pm 1} \Sigma_{x_2=0, \pm 1} \varphi(x_1, x_2) \cdot, \overline{\eta}(x_1, x_2) \quad (F.14)$$

と，ノルム $\| \varphi \| \equiv [(\varphi, \varphi)]$ 1/2 とが導入された可分なヒルベルト空間 $\mathfrak{H}$ において、

$$L = \{1, 2, \cdots, 25\} \quad (F.15)$$

を採用すれば、1次独立な系 $\{\psi_\ell\}_{\ell \in L}$ は次のように選んでよい：

文献

赤穂昭太郎，速水悟，長谷川修，吉村隆，麻生英樹："EM法を用いた複数情報源からの概念獲得"，電子情報通信学会論文誌A, vol.J80-A, no.9, pp.1546-1553, Sept.1997

に従って、次の1次独立な系 $\{\psi_\ell\}_{\ell \in L}$ を選定できる：

(1) $\psi_1(x_1, x_2) = \begin{cases} 1 \text{ if } \langle x_1, x_2 \rangle \in \{\langle 0, 0 \rangle\} \\ 0 \text{ otherwise} \end{cases}$

(2) $\psi_2(x_1, x_2) = \begin{cases} 1 \text{ if } \langle x_1, x_2 \rangle \in \{\langle 0, 0 \rangle, \langle 0, 1 \rangle\} \\ 0 \text{ otherwise} \end{cases}$

(3) $\psi_3(x_1, x_2) = \begin{cases} 1 \text{ if } \langle x_1, x_2 \rangle \in \{\langle 0, 0 \rangle, \langle 1, 1 \rangle\} \\ 0 \text{ otherwise} \end{cases}$

(4) $\psi_4(x_1, x_2) = \begin{cases} 1 \text{ if } \langle x_1, x_2 \rangle \in \{\langle 0, 0 \rangle, \langle 0, 1 \rangle\} \\ 0 \text{ otherwise} \end{cases}$

(5) $\psi_5(x_1, x_2) = \begin{cases} 1 \text{ if } \langle x_1, x_2 \rangle \in \{\langle 0, 0 \rangle, \langle -1, 1 \rangle\} \\ 0 \text{ otherwise} \end{cases}$

(6) $\psi_6(x_1, x_2) = \begin{cases} 1 \text{ if } \langle x_1, x_2 \rangle \in \{\langle 0, -1 \rangle, \langle 0, 0 \rangle, \langle 1, 0 \rangle\} \\ 0 \text{ otherwise} \end{cases}$

(7) $\psi_7(x_1, x_2) = \begin{cases} 1 \text{ if } \langle x_1, x_2 \rangle \in \{\langle -1, -1 \rangle, \langle 0, 0 \rangle, \langle 1, 1 \rangle\} \\ 0 \text{ otherwise} \end{cases}$

(8) $\psi_8(x_1, x_2) = \begin{cases} 1 \text{ if } \langle x_1, x_2 \rangle \in \{\langle 0, -1 \rangle, \langle 0, 0 \rangle, \langle 0, 1 \rangle\} \\ 0 \text{ otherwise} \end{cases}$

(9) $\psi_9(x_1, x_2) =$
$$\begin{cases} 1 \text{ if } \langle x_1, x_2 \rangle \in \{\langle -1, 1 \rangle, \langle 0, 0 \rangle, \langle 0, -1 \rangle\} \\ 0 \text{ otherwise} \end{cases}$$

(10) $\psi_{10}(x_1, x_2) =$
$$\begin{cases} 1 \text{ if } \langle x_1, x_2 \rangle \in \{\langle -1, 0 \rangle, \langle 0, 0 \rangle, \langle 1, 1 \rangle\} \\ 0 \text{ otherwise} \end{cases}$$

(11) $\psi_{11}(x_1, x_2) =$
$$\begin{cases} 1 \text{ if } \langle x_1, x_2 \rangle \in \{\langle -1, -1 \rangle, \langle 0, 0 \rangle, \langle 0, 1 \rangle\} \\ 0 \text{ otherwise} \end{cases}$$

(12) $\psi_{12}(x_1, x_2) =$
$$\begin{cases} 1 \text{ if } \langle x_1, x_2 \rangle \in \{\langle -1, 1 \rangle, \langle 0, 0 \rangle, \langle 0, -1 \rangle\} \\ 0 \text{ otherwise} \end{cases}$$

(13) $\psi_{13}(x_1, x_2) =$
$$\begin{cases} 1 \text{ if } \langle x_1, x_2 \rangle \in \{\langle -1, 0 \rangle, \langle 0, 0 \rangle, \langle 1, -1 \rangle\} \\ 0 \text{ otherwise} \end{cases}$$

(14) $\psi_{14}(x_1, x_2) =$
$$\begin{cases} 1 \text{ if } \langle x_1, x_2 \rangle \in \{\langle -1, -1 \rangle, \langle 0, 0 \rangle, \langle 1, 0 \rangle\} \\ 0 \text{ otherwise} \end{cases}$$

(15) $\psi_{15}(x_1, x_2) =$
$$\begin{cases} 1 \text{ if } \langle x_1, x_2 \rangle \in \{\langle 0, -1 \rangle, \langle 0, 0 \rangle, \langle 1, 1 \rangle\} \\ 0 \text{ otherwise} \end{cases}$$

(16) $\psi_{16}(x_1, x_2) =$
$$\begin{cases} 1 \text{ if } \langle x_1, x_2 \rangle \in \{\langle 0, 0 \rangle, \langle 0, 1 \rangle, \langle 1, -1 \rangle\} \\ 0 \text{ otherwise} \end{cases}$$

(17) $\psi_{17}(x_1, x_2) =$
$$\begin{cases} 1 \text{ if } \langle x_1, x_2 \rangle \in \{\langle -1, 1 \rangle, \langle 0, 0 \rangle, \langle 1, 0 \rangle\} \\ 0 \text{ otherwise} \end{cases}$$

(18) $\psi_{18}(x_1, x_2) =$
$$\begin{cases} 1 \text{ if } \langle x_1, x_2 \rangle \in \{\langle -1, 0 \rangle, \langle 0, 0 \rangle, \langle 0, 1 \rangle\} \\ 0 \text{ otherwise} \end{cases}$$

(19) $\psi_{19}(x_1, x_2) =$
$$\begin{cases} 1 \text{ if } \langle x_1, x_2 \rangle \in \{\langle -1, -1 \rangle, \langle -1, 1 \rangle, \langle 0, 0 \rangle\} \\ 0 \text{ otherwise} \end{cases}$$

(20) $\psi_{20}(x_1, x_2) =$
$$\begin{cases} 1 \text{ if } \langle x_1, x_2 \rangle \in \{\langle -1, 0 \rangle, \langle 0, 0 \rangle, \langle 0, -1 \rangle\} \\ 0 \text{ otherwise} \end{cases}$$

(21) $\psi_{21}(x_1, x_2) =$
$$\begin{cases} 1 \text{ if } \langle x_1, x_2 \rangle \in \{\langle -1, -1 \rangle, \langle 0, 0 \rangle, \langle 1, -1 \rangle\} \\ 0 \text{ otherwise} \end{cases}$$

(22) $\psi_{22}(x_1, x_2) =$
$$\begin{cases} 1 \text{ if } \langle x_1, x_2 \rangle \in \{\langle 0, -1 \rangle, \langle 0, 0 \rangle, \langle 1, 0 \rangle\} \\ 0 \text{ otherwise} \end{cases}$$

(23) $\psi_{23}(x_1, x_2) =$
$$\begin{cases} 1 \text{ if } \langle x_1, x_2 \rangle \in \{\langle 0, 0\rangle, \langle 1, 1\rangle, \langle 1, -1\rangle\} \\ 0 \text{ otherwise} \end{cases}$$
(24) $\psi_{24}(x_1, x_2) =$
$$\begin{cases} 1 \text{ if } \langle x_1, x_2 \rangle \in \{\langle 0, 0\rangle, \langle 0, 1\rangle, \langle 1, 0\rangle\} \\ 0 \text{ otherwise} \end{cases}$$
(25) $\psi_{25}(x_1, x_2) =$
$$\begin{cases} 1 \text{ if } \langle x_1, x_2 \rangle \in \{\langle -1, 1\rangle, \langle 0, 0\rangle, \langle 1, 1\rangle\} \\ 0 \text{ otherwise} \end{cases} \tag{F.16}$$
□

1次独立な系 $\{\psi_\ell\}_{\ell \in L}$ と、特徴抽出写像
$$u: \Phi \times L \to R \text{(実数全体の集合)} \tag{F.17}$$
とを選び、
$$T\varphi = \sum_{\ell \in L} u(\varphi, \ell) \cdot \psi_\ell \text{ for any } \varphi \in \Phi \tag{F.18}$$
で定義される "axiom 1の(i), (ii), (iii)の3後半と(iv)を満たす式(2.1)のモデル構成作用素" T を導入できる。ここに、$u(\varphi, \ell) \in R$ はパターン $\varphi \in \Phi$ から抽出された第 $\ell \in L$ 番目の特徴量である。式(F.18)の T については、φ から抽出された特徴量の組
$$\underline{u}(\varphi) \equiv \{u(\varphi, \ell) \mid \ell \in L\} \tag{F.19}$$
の保存性質
$$\forall \varphi \in \Phi, \underline{u}(T\varphi) = \underline{u}(\varphi) \tag{F.20}$$
が成立している。

式(F.18)の T を採用している場合、$\underline{u}(\varphi)$ を
$$\underline{u}(\varphi) \to \underline{u}(\eta) \tag{F.21}$$
と変換できた場合、パターンモデル $T\varphi$ を
$$T\varphi \to T\eta \tag{F.22}$$
というように、パターンモデル $T\eta$ へと変換できたことになるが、以後、式(2.1)のモデル構成作用素 T として、式(F.10)の T とは限らなくて、一般に、axiom 1を満たす対【Φ, T】での T が選ばれているとしよう。

F3. 簡単な単段階認識法

特徴軸の番号集合 L を、
$$L = \{1, 2, \cdots, n\} \tag{F.23}$$
として、2つの n 次元実ベクトル
$$\underline{a} = \mathrm{col}(a_1\ a_2 \cdots a_n) \text{(列ベクトル)}$$
$$\underline{b} = \mathrm{col}(b_1\ b_2 \cdots b_n) \tag{F.24}$$
間の内積
$$[\underline{a}, \underline{b}] = \sum_{\ell \in L} a_\ell \cdot b_\ell \tag{F.25}$$
と、ノルム
$$|\underline{a}| = \sqrt{(a, a)} \tag{F.26}$$
を導入する。

パターン$\varphi \in \Phi$ は、m 個のカテゴリ

$$\mathfrak{C}_1, \mathfrak{C}_2, \cdots, \mathfrak{C}_m \tag{F.27}$$

の内の1つ、例えば、第$j \in J$番目のカテゴリ$\mathfrak{C}_j$に分類するとしよう。ここに、カテゴリ番号jの集合

$$J=\{1, 2, \cdots, m\} \tag{F.28}$$

が導入され、全カテゴリ集合

$$\underline{\mathfrak{C}}=\{\mathfrak{C}_j \mid j \in J\} \tag{F.29}$$

も導入される。

その帰属するカテゴリが既知のパターン系列(訓練パターンの系列；a sequence of training patterns)

$$\varphi_{i,q}, q=1, 2, \cdots, n_i (i \in J) \tag{F.30}$$

を用意する。

$$z_{i,q}(\ell)=u(\varphi_{i,q}, \ell), \ \ell \in L \tag{F.31}$$

とおく。カテゴリ$\mathfrak{C}_j$の、第$\ell \in L$番目の特徴量の平均値

$$z_{mean}(i ; \ell)=(1/n_i)\cdot \sum_{q=1}^{n_i} z_{i,q}(\ell) \tag{F.32}$$

を定義し、これを第$\ell \in L$番目の成分とするベクトル(平均特徴ベクトル)

$$\underline{z_{mean}}(i) =col(z_{mean}(i ; 1)\ z_{mean}(i ; 2), \cdots, \ z_{mean}(i ; n)) \tag{F.33}$$

を設ける。処理の対象とする問題のパターン$\varphi \in \Phi$について、特徴量の組$\underline{u}(\varphi)$を算出し、

$$z(\ell)=u(\varphi, \ell) \tag{F.34}$$

を第$\ell \in L$番目の成分とする特徴ベクトル

$$\underline{z}=col(z_1\ z_2 \cdots z_n) \tag{F.35}$$

のノルム規格化ベクトル$z|z|^{-1}$と、平均特徴ベクトルのノルム規格化ベクトル$\underline{z_{mean}}(i)|\underline{z_{mean}}(i)|^{-1}$との距離

$$|\underline{z}|\underline{z}|^{-1}-\underline{z_{mean}}(i)|\underline{z_{mean}}(i)|^{-1}| \tag{F.36}$$

を各$i \in J$にわたり求め、その距離が最小となるカテゴリ番号

$$\arg\min_{i \in J} |\underline{z}|\underline{z}|^{-1}- \underline{z_{mean}}(i)|\underline{z_{mean}}(i)|^{-1}| =j \in J \tag{F.37}$$

を見つけ、

パターンφは第$j \in J$番目のカテゴリ$\mathfrak{C}_j$に帰属する (F.38)

と分類(認識)する。

$$|\underline{z}|\underline{z}|^{-1}-\underline{z_{mean}}(i)|\underline{z_{mean}}(i)|^{-1}|^2 =2\cdot[1-[\underline{z}, \underline{z_{mean}}(i)] /\{|\underline{z}|^{-1}\cdot|\underline{z_{mean}}(i)|^{-1}] \tag{F.39}$$

であるから、その規格化内積

$$[\underline{z}, \underline{z_{mean}}(i)] /\{|\underline{z}|^{-1}\cdot|\underline{z_{mean}}(i)|^{-1}\} \tag{F.40}$$

が最大となるカテゴリ番号

$$\arg\max_{i \in J}\ [\underline{z}, \underline{z_{mean}}(i)]$$

$/\{|\underline{z}|^{-1}\cdot|\underline{z_{mean}}(i)|^{-1}\}\cdot$

$= j\in J$ (F.41)

を見つけ、式(F.38)の如く、分類(認識)するとしてよい。

F4. 特徴量を一致させ、認識する手法

前章の分類法を少し、改良しよう。処理の対象とする問題のパターン $\varphi\in\Phi$ が抽出される特徴量 $\underline{u}(\varphi)$ に関し少しぐらい変動していても、つまり、φ が少しぐらい変形していても、良好な認識結果が得られるように改良しよう。

F4.1 想起形認識への転換

各カテゴリ $\mathfrak{C}_i(i\in J)$ にわたり行列 A_i を用意し、カテゴリ番号

$\arg\min\ |A_i\underline{z}-\underline{z_{mean}}(i)|$

$= j\in J$ (F.42)

を求め、

パターン φ の特徴量は $A_i\underline{z}$ として再現され、

φ は第 $j\in J$ 番目のカテゴリ $\mathfrak{C}_j$ に帰属する (F.43)

と分類(認識)する。

もし、モデル構成作用素 T として、式(F.18)の T を採用しているならば、式(F.43)は次のように、想起形(連想形)認識の形式に書き替えられてよい：

式(F.34)の如く設定された特徴量の組

$\underline{z}=\underline{u}(\varphi)$ (F.44)

を持つようなパターンモデル $T\varphi$ を見たり聞いたりしたならば、パターン φ かのように見えたり聞こえたりするパターン φ は、

$\underline{u}(\eta)=A_i\underline{z}$ (F.45)

として、パターン

$T\eta=\sum_{\ell\in L}u(\eta,\ell)\cdot\psi_\ell$ (F.46)

として再現され(パターン想起の働き)、

φ は第 $j\in J$ 番目のカテゴリ $\mathfrak{C}_j$ に帰属する(分類・認識の働き)。 (F.47)

□

F4.2 特徴変換行列 A_i の設定法

残されている問題は、特徴変換行列 A_i をどう設定するかである。

同じカテゴリ $\mathfrak{C}_i$ に帰属する式(F.30)の訓練パターン $\varphi_{i,q}$, $q=1, 2, \cdots, n_i$ の特徴量の組

$\underline{u}(\varphi_{i,q}),\ q=1, 2, \cdots, n_i$ (F.48)

をすべて、あらかじめわかっている特徴量

$\underline{v}(i)=\mathrm{col}(v(i)_1\ v(i)_2\cdots v(i)_n)$ (F.49)

に

$A_i\underline{z_{i,q}}=\underline{v}(i),\ q=1, 2, \cdots, n_i$ (F.50)

というように変換する行列 A_i を見つければよい。ここに、

式(F.33)の平均特徴ベクトル zmean(i)に注目し、

$\underline{v}(i)=\underline{z_{mean}}(i)$ (F.51)

としてもよいが、理想出力 $\underline{v}(i)$ 間に分離条件

$\forall j \in J,\ \forall i \in J-\{i\},\ \underline{v}(i) \neq \underline{v}(j)$ (F.52)

を満たしている必要がある。

式(F.49)の$\underline{v}(i)$はカテゴリ$\mathfrak{C}_i$の持つ諸性質を典型的に代表する代表パターンω_iから抽出される特徴量の組$\underline{u}(\omega_i)$であると考えられ、等式(F.50)の成立は望ましいことであるが、一般的には不可能である。そこで、

「すべての$\underline{z_{i,q}}$, $q=1, 2, \cdots, n_i$を$\underline{v}(i)$に近ければ近い程、
$\underline{v}(i)$の近傍により近く変換するような特徴変換行列A_iを
見い出す」 (F.53)

ことにしよう。$A_i\ \underline{z_{i,q}}$の$\underline{v}(i)$からの自乗変換誤差を

$\varepsilon_{i,q} \equiv |A_i\ \underline{z_{i,q}} - \underline{v}(i)|^2$ (F.54)

と表せば、第$i \in J$番目のカテゴリ$\mathfrak{C}_i$の各訓練パターン$\varphi_{i,q}$に対して、平均自乗変換誤差ε_iは、

$\varepsilon_i = n_i^{-1} \cdot \sum_{q=1}^{n_i} \varepsilon_{i,q}$ (F.55)

となる。このε_iを最小とするように行列A_iを決定すればよい。A_iは$a(i)_{k\ell}$を第$k \in L$行第$\ell \in L$列とする$|L| \times |L|$の正方行列であり、

$A_i = (a(i)_{k\ell})k,\ \ell \in L$ (F.56)

と表されるが、次の定理F.1が成立し、ε_iを最小ならしめる解行列A_iが求まる。

[定理F.1] **(平均自乗変換誤差 ε_i の解行列 A_i の定理)**

平均自乗変換誤差ε_iを最小ならしめる解行列A_iは、

A_i
$= [n_i^{-1} \cdot \sum_{r=1}^{n_i} \underline{v}(i) \cdot \underline{z_{i,r}}^t] \cdot$
$[n_i^{-1} \cdot \sum_{s=1}^{n_i} \underline{z_{i,s}} \cdot \underline{z_{i,s}}^t]^{-1}$ (F.57)

と求められる。$\underline{z_{i,q}}^t$は$\underline{z_{i,q}}$の転置ベクトルであり、行ベクトルである。 □

先ず、式(F.23)の集合Lについて、

$A = (a_{ij})_{i,j \in L} = (a_{ij})_{1 \leqq i,j \leqq n}$のとき、
$\partial / \partial A = (\partial / \partial a_{ij})_{1 \leqq i,j \leqq n}$ (F.58)

の意であることに注意して、次の2補助定理F.1, , F.2を証明する。

[補助定理F.1] (双1次形式 [Ay, x] の微分)

$(\partial / \partial A)\ [A\underline{y}, \underline{x}]$
$= \underline{x} \cdot \underline{y}^t.$ (F.59)

(証明) $(\partial / \partial a_{ij})\ [A\underline{y}, \underline{x}]$
$= (\partial / \partial a_{ij})\ [x_i \cdot \sum_{\ell \in L} a_{i\ell} \cdot y_\ell$
$+ \sum_{k \in L-\{i\}} x_k \cdot \sum_{\ell \in L} a_{k\ell} \cdot y_\ell]$
$= (\partial / \partial a_{ij})\ [x_i \cdot \{a_{ij} \cdot y_j + \sum_{\ell \in L-\{j\}} a_{i\ell} \cdot y_\ell\}$
$+ \sum_{k \in L-\{i\}} x_k \cdot \sum_{\ell \in L} a_{k\ell} \cdot y_\ell]$
$= x_i \cdot y_j$

を得、

$(\partial / \partial A)\ [A\underline{y}, \underline{x}]$
$= (x_i \cdot y_j)\ 1 \leqq i, j \leqq n$
$= \underline{x} \cdot \underline{y}^t.$ □

［補助定理F.2］（2次形式［$A\underline{x}$，$\underline{x}$］の微分）

$$(\partial/\partial A)\ [A\underline{x},\ \underline{x}]$$
$$=2\cdot A\underline{x}\cdot\underline{x}^{t}. \qquad \text{(F.60)}$$

（証明）$(\partial/\partial a_{ij})\ [A\underline{x},\ \underline{x}]$

$$=(\partial/\partial a_{ij})\sum_{k\in L}\{a_{kj}\cdot x_j+\sum_{\ell\in L-\{j\}} a_{k\ell}\cdot x_\ell\}\cdot\{a_{kj}\cdot x_j+\sum_{q\in L-\{j\}} a_{kq}\cdot x_q\}$$
$$=(\partial/\partial a_{ij})\sum_{k\in L}\Big[a_{kj}^2\cdot x_j^2+a_{kj}\cdot x_j\cdot\{\sum_{q\in L-\{j\}} a_{kq}\cdot x_q\}$$
$$+\{\sum_{\ell\in L-\{j\}} a_{k\ell}\cdot x_\ell\}\cdot a_{kj}\cdot x_j$$
$$+\{\sum_{\ell\in L-\{j\}} a_{k\ell}\cdot x_\ell\}\cdot\{\sum_{q\in L-\{j\}} a_{iq}\cdot x_q\}\Big]$$
$$=2a_{ij}\cdot x_j^2+x_j\cdot\{\sum_{q\in L-\{j\}} a_{iq}\cdot x_q\}$$
$$+\{\sum_{\ell\in L-\{j\}} a_{i\ell}\cdot x_\ell\}\cdot x_j$$
$$=x_j\cdot\{\sum_{q\in L} a_{iq}\cdot x_q\}+\{\sum_{\ell\in L} a_{i\ell}\cdot x_\ell\}\cdot x_j$$
$$=2\cdot\{\sum_{\ell\in L} a_{i\ell}\cdot x_\ell\}\cdot x_j$$

であるから、

$$(\partial/\partial A)[A\underline{x},\ \underline{x}]$$
$$=2\cdot A\underline{x}\cdot\underline{x}^{t}.$$

□

（定理F.1の証明）　以下に、式(F.57)を導出する。

先ず、式(F.55)の ε_i を変形すれば、

$$\varepsilon_i$$
$$=n_i^{-1}\cdot\sum_{q=1}^{n_i}\ [A_i\underline{z_{i,q}}-\underline{v}(i),\ A_i\underline{z_{i,q}}-\underline{v}(i)] \qquad \because\ 式(F.54)$$
$$=n_i^{-1}\cdot\sum_{q=1}^{n_i}\Big[[A_i\underline{z_{i,q}},\ A_i\underline{z_{i,q}}]$$
$$-2\cdot[A_i\underline{z_{i,q}},\ \underline{v}(i)]-[\underline{v}(i),\ \underline{v}(i)]\Big] \qquad \text{(F.61)}$$

であるから、2補助定理F.1，F.2を適用して、

$$0=(\partial\varepsilon_i/\partial A)$$
$$=n_i^{-1}\cdot\sum_{q=1}^{n_i}\Big[2\cdot A_i\underline{z_{i,q}}\cdot\underline{z_{i,q}}^{t}-2\cdot\underline{v}(i)\cdot\underline{z_{i,q}}^{t}\Big]$$
$$\therefore\ n_i^{-1}\cdot\sum_{q=1}^{n_i} A_i\underline{z_{i,q}}\cdot\underline{z_{i,q}}^{t}$$
$$=n_i^{-1}\cdot\sum_{q=1}^{n_i}\ \underline{v}(i)\cdot\underline{z_{i,q}}^{t} \qquad \text{(F.62)}$$

を得、方程式(F.62)の解 A_i は、式(F.57)で与えられることを示せば、証明が終わる。

実際、

式(F.62)の左辺

$$=n_i^{-1}\cdot\sum_{q=1}^{n_i}\Big[[n_i^{-1}\cdot\sum_{r=1}^{n_i}\ \underline{v}(i)\cdot\underline{z_{i,r}}^{t}]\cdot[n_i^{-1}\cdot\sum_{s=1}^{n_i}\ \underline{z_{i,s}}\cdot\underline{z_{i,s}}^{t}]^{-1}\underline{z_{i,q}}\Big]\cdot\underline{z_{i,q}}^{t} \qquad \because\ 式(F.57)$$
$$=\ [n_i^{-1}\cdot\sum_{r=1}^{n_i}\ \underline{v}(i)\cdot\underline{z_{i,r}}^{t}]\cdot[n_i^{-1}\cdot\sum_{s=1}^{n_i}\ \underline{z_{i,s}}\cdot\underline{z_{i,s}}^{t}]^{-1}\cdot[n_i^{-1}\cdot\sum_{q=1}^{n_i}\ \underline{z_{i,q}}\cdot\underline{z_{i,q}}^{t}]$$

$= [n_i^{-1} \cdot \sum_{r=1}^{n_i} \underline{y}(i) \cdot \underline{z_{i,r}}^t] \cdot$

=式(F.62)の右辺 (F.63)

を得、行列 A_i は方程式(F.62)を満たす。 □

尚、定理F.1の証明は、基本的には、文献

濱裕光，柳重堪，阮牧："シグナルとシステムの数学"，森北出版，pp.128-131, Dec.1997 に基づく。

付録G. 整形化パターンモデル構成作用素 T

本付録Gでは、処理の対象とする問題の、変形しているパターン φ を1次展開係の組 $a(\varphi)$ を整形化して得られるパターンモデル $T\varphi$ を研究する。

G.1 モデル構成作用素 T の構成

定理2.1を適用できるような式(2.1)の写像 T を構成する。

式(F.23)の集合 L を採用し、3定義式(F.24)～(F.26)の下で、論じる。

式(F.9)を最小ならしめる各1次結合係数 $a_\ell(\varphi)$ $(\ell \in L)$ の組

$$\underline{a}(\varphi) = \mathrm{col}(a_1(\varphi) a_2(\varphi) \cdots a_n(\varphi)) \tag{G.1}$$

を導入する。φ は処理の対象とする問題のパターンである。φ の集まり Φ は可分なあるヒルベルト空間 $\mathfrak{H}$ の、零元 0 を含むある部分集合である。

大きさ $n(=|L|)\times n$ の実行列 B_i と、実振幅係数 c_i を導入し、$\underline{a}(\varphi)$ を変換して得られるベクトル $c_i \cdot B_i \underline{a}(\varphi)$ を第$i \in J$番目のカテゴリ $\mathfrak{C}_i$ の代表パターン ω_i の 1 次結合係数組 $\underline{a}(\omega_i)$ に近づけることを考え、汎関数

$$F_i(\varphi) \equiv 2^{-1} \cdot | c_i \cdot B_i \underline{a}(\varphi) - \underline{a}(\omega_i) |^2 \tag{G.2}$$

を導入する。$F_i(\varphi)$ を最小ならしめる振幅 c_i と、特徴変換行列と呼ばれてよい B_i を各々、

$$c_i(\varphi), \; B_i(\varphi) \tag{G.3}$$

と表す。固定したパターン $\varphi \in \Phi$ について、汎関数 $F_i(\varphi)$ の最小値 $\min_{i \in J} F_i(\varphi)$ を与えるカテゴリ番号

$$\arg \min_{i \in J} F_i(\varphi) = j \in J \tag{G.4}$$

を求める。得られたベクトル $B_j(\varphi) \underline{a}(\varphi)$ の第 $\ell \in L$ 番目の成分を $(B_j(\varphi) \underline{a}(\varphi))_\ell$ と表すと、式(F.12)のように展開される原パターン φ に対応して、

パターン

$$\eta = c_j(\varphi) \cdot \sum_{k \in L} (B_j(\varphi) \underline{a}(\varphi))_k \cdot \psi_k \tag{G.5}$$

が存在することになるが、更に、$B_j(\varphi) \underline{a}(\varphi)$ を規格化して、パターン

$$T\varphi = \begin{cases} c_j(\varphi) \cdot \sum_{\ell \in L} \left[(B_j(\varphi) \underline{a}(\varphi))_\ell \right. \\ \quad \left. / \sup_{k \in L} | (B_j(\varphi) \underline{a}(\varphi))_k | \right] \cdot \psi_\ell \\ \quad \cdots \forall k \in L, \; (B_j(\varphi) \underline{a}(\varphi))_k > 0 \text{ の場合} \\ 0 \quad \cdots \exists k \in L, \; (B_j(\varphi) \underline{a}(\varphi))_k = 0 \text{ の場合} \end{cases} \tag{G.6}$$

を定義する。このとき、次の定理G.1が成立し、φ に対応して、パターンモデル $T\varphi$ が構成されたことがわかる。

［定理G.1］（**各1次結合係数 $a_\ell(\varphi)(\ell\in L)$ の整形化に基づくパターンモデル構成定理**）

2条件

$$\forall j\in J,\ \exists \ell\in L,\ a_\ell(\omega_j)\neq 0 \tag{G.7}$$

$$\forall j\in J,\ \forall i\in J-\{j\},\ \exists \ell\in L,$$
$$a_\ell(\omega_i)\neq a_\ell(\omega_j) \tag{G.8}$$

の下で、式(G.6)のごとく定義された式の写像 T は、axiom 1の(i)，(ii)，(iii)の3後半、並びに、(iv)を満たす。

［定理G.1の系1］

$$\forall \ell\in L,\ a_\ell(\varphi)=0 \tag{G.9}$$

であれば、

$$\varphi=\varphi_\perp \text{ であり、} T\varphi=0. \tag{G.10}$$

（証明）系1の成立を示そう。式(F.12)より、$\varphi=\varphi_\perp$ が成立し、

$$\forall i\in J,\ B_i(\varphi)\underline{a}(\varphi)=0 \tag{G.11}$$

を得、T の定義式(G.6)より、$T\varphi=0$ が成り立つ。

axiom 1の(i)の後半の成立：　$\varphi=0$ とすれば、$\varphi_\perp=0$ が成立し、式(G.9)が成立し、系1より、(i)の後半が成立する。

axiom 1の(ii)の後半の成立：　b を正定数とし、

$$\eta=b\cdot\varphi \tag{G.12}$$

とする。

（イ）$\varphi=\varphi_\perp$ のとき

系1より、$T\varphi=0$ である。

$$\forall \ell\in L,\ a_\ell(\eta)=b\cdot a_\ell(\varphi) \tag{G.13}$$
$$=0 \quad \because \text{ 式(G.9)}$$

を得、系1より、$T\eta=0$ である。

$$T\eta=T\varphi \tag{G.14}$$
$$=0 \tag{G.15}$$

（ロ）$\varphi\neq\varphi_\perp$ のとき

このとき、式(F.12)より、

$$\exists \ell\in L,\ a_\ell(\varphi)\neq 0 \tag{G.16}$$

である。

$$\min_{i\in J} F_i(\varphi)=F_j(\varphi) \tag{G.17}$$

とする。式(G.13)が成り立ち、

$$c_j(\eta)=c_j(\varphi) \tag{G.18}$$
$$B_j(\eta)=B_j(\varphi) \tag{G.19}$$

ととれば、式(G.13)が成り立っているから、

$$(B_j(\eta)\underline{a}(\eta))_\ell/\sup_{\ell\in L}|(B_j(\eta)\underline{a}(\eta))_\ell|$$
$$=(b/b)\cdot(B_j(\varphi)a(\varphi))\ \ell$$
$$/\sup_{\ell\in L}|(B_j(\varphi)\underline{a}(\varphi))_\ell|$$

$=(B_j(\varphi)\underline{a}(\varphi))_\ell$
$/\sup_{\ell \in L} |(B_j(\varphi)\underline{a}(\varphi))_\ell|$ (G.20)
$\neq \underline{0}$ (G.21)

を得、T の定義式(G.6)より、式(G.14)が成り立つ。

axiom 1の(iii)の後半の成立：

$\eta = T\varphi$ (G.22)

とおく。

(ロ) $\eta=0$ の場合

系1より、$T\eta=0$ を得、2式(G.14)，(G.15)が成り立つ。

(ハ) $\eta\neq 0$ の場合

式(G.6)より、

$\eta = T\varphi =$
$c_j(\varphi)\cdot\sum_{\ell \in L}[(B_j(\varphi)\underline{a}(\varphi))_\ell$
$/\sup_{k \in L}|(B_j(\varphi)\underline{a}(\varphi))_k|]\cdot\psi_\ell$ (G.23)

である。明らかに、

$c_j(\varphi)\neq 0$ (G.24)

である。式(F.10)より、

$\forall \ell \in L,\ a_\ell(\eta)$
$=c_j(\varphi)\cdot[(B_j(\varphi)\underline{a}(\varphi))\ell$
$/\sup_{k \in L}|(B_j(\varphi)\underline{a}(\varphi))_k|]$ (G.25)

が成立している。更に、

$\eta_\perp=0 \quad \because \quad$ 2式(G.6)，(F.12) (G.26)

である。

各 $F_i(\eta)$を最小にする $c_i(\eta)$，$B_i(\eta)$を求めたとしよう。式(G.16)が成立したとすれば、

$\min_{i \in J} F_i(\eta)=F_j(\eta)$ (G.27)

を満たす $c_j(\eta)$，$B_j(\eta)$は、

$c_j(\eta)=1$ (G.28)
$B_j(\eta)$
$=c_j(\varphi)^{-1}\cdot\sup_{k \in L}|(B_j(\varphi)\underline{a}(\varphi))_k|]\cdot B_j(\varphi)$ (G.29)

であり、

$T\eta$
$=c_j(\eta)\cdot\sum_{\ell \in L}[(B_j(\eta)\underline{a}(\eta))_\ell$
$/\sup_{k \in L}|(B_j(\eta)\underline{a}(\eta))_k|]\cdot\psi_\ell$
$=\sum_{\ell \in L}[(B_j(\eta)\underline{a}(\eta))_\ell$
$/\sup_{k \in L}|(B_j(\eta)\underline{a}(\eta))_k|]\cdot\psi_\ell$
$=\sum_{\ell \in L}(B_j(\varphi)\underline{a}(\varphi))_\ell$
$/\sup_{k \in L}|(B_j(\varphi)\underline{a}(\varphi))_\ell]\cdot\psi_\ell$
$\because$ 2式(G.24)，(G.25) (G.30)

$= T\varphi$ (G.31)

$= \eta$ (G.32)

を得、式(G.14)が成立する。

axiom 1の(iv)の成立：

$\varphi = \omega_j$とすれば、

$c_j(\varphi) = 1$ (G.33)

$B_j(\varphi) = I$(単位行列) (G.34)

のとき、式(G.16)が

$F_j(\varphi) = 0$ (G.35)

を得、式(G.16)が成り立つ。よって、2条件式(G.7), (G.8)を考慮すれば、

$$T\varphi = \sum_{\ell \in L} [a_\ell(\varphi) / \sup_{k \in L} | a_\ell(\varphi) |] \cdot \psi_\ell$$

$$(= \sup_{k \in L} | a_\ell(\varphi) | \cdot [\varphi - \varphi_\perp] \quad \because \text{ 式(F.12)}) \quad \text{(G.36)}$$

$$\neq 0 \quad \text{(G.37)}$$

が得られる。 □

G.2 汎関数 $F_i(\varphi)$を最小ならしめる振幅 $c_i(\varphi)$と、各1次結合係数 $a_\ell(\varphi)$の変換行列 B_i の逐次的決定

式(G.2)の汎関数 $F_i(\varphi)$を最小ならしめるように、振幅 c_i と、各 1 次結合係数 $a_\ell(\varphi)$の変換行列 B_i とを逐次的に決定し、$c_i(\varphi)$と $B_i(\varphi)$とを得るために、最急降下法を適用しよう。

G2.1 振幅 $c_i(\varphi)$の決定

式(G.2)の汎関数 $F_i(\varphi)$を最小ならしめる振幅 $c_i(\varphi) = c_i$ を求めよう。

式(G.2)の $F_i(\varphi)$は、

$$F_i(\varphi) = 2^{-1} \cdot \sum_{\ell \in L} [c_i \cdot (B_i \underline{a}(\varphi))_\ell - a_\ell(\omega_i)]^2 \quad \text{(G.38)}$$

と表現されることに注意し、

$$\partial F_i(\varphi) / \partial c_i = \sum_{\ell \in L} [c_i \cdot (B_i \underline{a}(\varphi))_\ell - a_\ell(\omega_i)] \cdot (B_i \underline{a}(\varphi))_\ell \quad \text{(G.39)}$$

$$= 0 \quad \text{(G.40)}$$

を c_i に関し解けば、

$$c_i(\varphi) = [\underline{a}(\omega_i), B_i \underline{a}(\varphi)] / [B_i \underline{a}(\varphi), B_i \underline{a}(\varphi)] \quad \text{(G.41)}$$

と求まる。

G2.2 特徴変換行列 $B_i(\varphi)$の決定

式(G.2)の汎関数 $F_i(\varphi)$を最小ならしめる特徴変換行列 $B_i = B_i(\varphi)$を求めよう。

$$B_i = (b_{\ell, k}(i))_{\ell, k \in L} \quad \text{(G.42)}$$

とおけば、式(G.38)内の $(B_i \underline{a}(\varphi))_q$ は

$(B_i \underline{a}(\varphi))_q$

$$=\sum_{r\in L} b_{qr}(i)\cdot a_r(\varphi)$$
$$=b_{q,k(i)}\ a_k(\varphi)+\sum_{r\in L-\{k\}} b_{q,r}(i)\cdot a_r(\varphi) \qquad \text{(G.43)}$$
と分解できるから、
$$\delta_{\ell q}=1 \text{ if } \ell=q,\ =0 \text{ if } \ell\neq q \qquad \text{(G.44)}$$
として、
$$0=\partial F_i(\varphi)/\partial b_{\ell,k}(i)$$
$$=\sum_{q\in L}[c_i\cdot(B_i\,\underline{a}(\varphi))_q-a_q(\omega_i)]\cdot$$
$$c_i\cdot a_k(\varphi)\cdot\delta_{\ell q} \qquad \text{(G.45)}$$
$$=[c_i\cdot(B_i\,\underline{a}(\varphi))_\ell-a_\ell(\omega_i)]\cdot c_i\cdot a_k(\varphi)$$
$$=[c_i\cdot b_{\ell,k}(i)a_k(\varphi)+c_i\cdot\sum_{r\in L-\{k\}} b_{\ell,r}(i)$$
$$\cdot a_r(\varphi)-a_\ell(\omega_i)]\cdot c_i\cdot a_k(\varphi)$$
$$=b_{\ell,k}(i)\cdot c_i^2\cdot a_k(\varphi)^2+c_i^2\cdot a_k(\varphi)\cdot$$
$$\sum_{r\in L-\{k\}} b_{\ell,r}(i)\cdot a_r(\varphi)$$
$$-c_i\cdot a_k(\varphi)\cdot a_\ell(\omega_i) \qquad \text{(G.46)}$$
であり、$b_{\ell,k}(i)$に関し解けば、
$$b_{\ell,k}(i)$$
$$=[c_i^2\cdot a_k(\varphi)^2]^{-1}\cdot【c_i\cdot a_k(\varphi)\cdot a_\ell(\omega_i)$$
$$-c_i^2\cdot a_k(\varphi)\cdot\sum_{r\in L-\{k\}} b_{\ell,r}(i)\cdot a_r(\varphi)】 \qquad \text{(G.47)}$$
$$=[c_i\cdot a_k(\varphi)]^{-1}\cdot a_\ell(\omega_i)$$
$$-a_k(\varphi)^{-1}\cdot\sum_{r\in L-\{k\}} b_{\ell,r}(i)\cdot a_r(\varphi) \qquad \text{(G.48)}$$
と求まる。

G2.3　最急降下法による特徴変換行列 B_i の第 ℓ 行第 k 列成分 $b_{\ell,k}(i)$ の逐次的決定

最急降下法(学習法)によって、特徴変換行列 B_i の第 ℓ 行第 k 列成分 $b_{\ell,k}(i)$ を逐次的に決定しよう。解析的には、$b_{\ell,k}(i)$ は式(G.48)のごとく求められていることに注意しておく。

振幅 c_i が $c_i(\varphi)$ と求まっているとして、式(G.2)の汎関数 $F_i(\varphi)$ を最小ならしめる特徴変換行列 $B_i(\varphi)$ の $b_{\ell,k}(i)$ を最急降下方程式(G.52)を解くことにより、求めよう。

学習時刻変数 $s(\geqq 0)$ と、
$$\tau_{\ell,k}(i\,;\,s)\to 0(s\to+\infty) \qquad \text{(G.49)}$$
を満たす正値減少関数 $\tau_{\ell,k}(i\,;\,s)$ を導入し、式(G.42)の B_i を
$$B_i(\varphi\,;\,s)=(b_{\ell,k}(i)(\varphi\,;\,s))_{\ell,k\in L} \qquad \text{(G.50)}$$
と表し、式(G.38)の $F_i(\varphi)$ を
$$F_i(\varphi\,;\,s)$$
$$=2^{-1}\cdot\sum_{\ell\in L}[c_i\cdot(B_i(\varphi\,;\,s)\underline{a}(\varphi))_\ell-a_\ell(\omega_i)]^2 \qquad \text{(G.51)}$$
と表し、最急降下方程式
$$db_{\ell,k}(i\,;\,s)/ds$$
$$=-\tau_{\ell,k}(i\,;\,s)\cdot\partial F_i(\varphi\,;\,s)/\partial b_{\ell,k}(i\,;\,s)$$
$$,\ 0\leqq s<+\infty,\ \ \ell,\ k\in L,\ \ i\in J \qquad \text{(G.52)}$$
を考えると、
$$dF_i(\varphi\,;\,s)/ds$$
$$=\sum_{\ell\in L}\sum_{k\in L}[\partial F_i(\varphi\,;\,s)/\partial b_{\ell,k}(i\,;\,s)]$$

$\cdot[db_{\ell,k}(i;s)/ds]$

$$=-\sum_{\ell\in L}\sum_{k\in L}\tau_{\ell,k}(i;s)\cdot[\partial F_i(\varphi;s)/\partial b_{\ell,k}(i;s)]^2 \leqq 0 \qquad \text{(G.53)}$$

が成り立ち、汎関数 $F_i(\varphi;s)$ が減少するように、$b_{\ell,k}(i;s)$ の時間的発展が得られている。

正値減少関数 $\tau_{\ell,k}(i;s)$ は、

$$\tau_{\ell,k}(i;s)=1/[s+1]>0 \qquad \text{(G.54)}$$

と選ぶのがよい。

最急降下方程式(G.52)を離散時刻変数 $s=0, 1, 2, \cdots$ について書き直そう。

$$b_{\ell,k}(i;s+1) = b_{\ell,k}(i;s)+\Delta b_{\ell,k}(i;s),\quad s=0,1,2,\cdots \qquad \text{(G.55)}$$

を設ける。ここに、増分 $\Delta b_{\ell,k}(i;s)$ は、最急降下方程式(G.52)から、

$$\Delta b_{\ell,k}(i;s) = -\tau_{\ell,k}(i;s)\cdot\partial F_i(\varphi;s)/\partial b_{\ell,k}(i;s) \qquad \text{(G.56)}$$

と与えればよい。式(G.56)の $\Delta b_{\ell,k}(i;s)$ 内の　は、式(G.46)から、

$$\begin{aligned}&\partial F_i(\varphi;s)/\partial b_{\ell,k}(i;s)\\&=b_{\ell,k}(i;s)\cdot c_i(\varphi)^2\cdot a_k(\varphi)^2\\&+c_i(\varphi)^2\cdot a_k(\varphi)\cdot\sum_{r\in L-\{k\}} b_{\ell,r}(i;s)\\&\cdot a_r(\varphi)\\&-c_i(\varphi)\cdot a_k(\varphi)\cdot a_\ell(\omega_i)\end{aligned} \qquad \text{(G.57)}$$

と与えられる。

G.3　$c_i(\varphi)$ と $B_i(\varphi)$ との、同時学習の終了判定

実際には、式(G.2)の汎関数 $F_i(\varphi)$ を最小ならしめる c_i, B_i を同時に求めなければならない。この同時学習は、式(G.48)を利用しない形式で、次のように実行すればよい。

離散学習時刻 $t(=0, 1, 2, \cdots)$ を設ける。

$t=0$ のとき、

$$B_{i,t}\mid_{t=0}=I \qquad \text{(G.58)}$$

として、式(G.41)から、

$$\begin{aligned}&c_{i,t}(\varphi)\mid_{t=0}\\&=[\underline{a}(\omega_i),\ B_{i,t}\mid_{t=0}\underline{a}(\varphi)]\\&/[B_{i,t}\mid_{t=0}\underline{a}(\varphi),\ B_{i,t}\mid_{t=0}\underline{a}(\varphi)]\end{aligned} \qquad \text{(G.59)}$$

と求め、t を 0 に固定して、その後、G2.3節の最急降下法で、各 $b_{\ell,k,t}(i)\mid_{t=0}$ を求める。

$t=1, 2, 3, \cdots$ についても同様であり、$t(=1, 2, \cdots)$ を固定して、$B_{i,t-1}$ を使って、式(G.41)から、$c_{i,t}(\varphi)$ を

$$\begin{aligned}&c_{i,t}(\varphi)\\&=[\underline{a}(\omega_i),\ B_{i,t-1}\underline{a}(\varphi)]\\&/[B_{i,t-1}\underline{a}(\varphi),\ B_{i,t-1}\underline{a}(\varphi)]\end{aligned} \qquad \text{(G.60)}$$

と求めた後、t を固定して、G2.3節の最急降下法で、$s\to\infty$ として、各 $b_{\ell,k}(i;t)$ を求める。

上述の、 c_i, B_i の同時学習が終了することを判定するには、あらかじめ、2つの正数 $\varepsilon_1(t)$, $\varepsilon_2(t)$ を決定しておいて、不等式

$$①\sum_{i \in J} \mid c_{i,t+1}(\varphi) - c_{i,t}(\varphi) \mid < \varepsilon_1(t) \tag{G.61}$$

$$②\sum_{i \in J} \sum_{\ell \in L} \sum_{k \in L} \mid b_{\ell,k}(i;t+1) - b_{\ell,k}(i;t) \mid < \varepsilon_2(t) \tag{G.62}$$

を満たす時刻 t を決定し、

$$c_i(\varphi) = c_{i,t}(\varphi) \tag{G.63}$$

$$b_{\ell,k}(i;\varphi) = b_{\ell,k}(i;t), \quad \ell, k \in L \tag{G.64}$$

と求めればよい。ここに、

$$B_i(\varphi) = (b_{\ell,k}(i;\varphi))_{\ell,k \in L} \tag{G.65}$$

である。

上述をすべてのカテゴリ番号 $i \in J$ にわたって実行すればよい。

（鈴木昇一，文教大学・情報学部・情報システム学科，“文教大学・情報学部・情報研究 no.26” 投稿論文，論文題目 Support Vector Machine を利用した大分類関数の構成，投稿年月日 2001年8月13日（月））

（完了）