数理解析研究所講究録
975巻 1996年 145-154

パタン認識における不変則

安田道夫 Michio YASUDA*

Abstract

学習と識別は、パタン認識の必須の要素である。また、学習のためには、記憶が不可欠の要素である。したがって、パタン認識における学習・記憶・識別の三つの要素の性質と、相互の関連について考察しておくことは、パタン認識の研究を進めるために有意義であると同時に、必要な作業である。

1 はじめに

人のパタン認識(Pattern Recognition)機能とは、感覚器官で観測した、多様な外界の事象を、複数の既知のカテゴリ概念の何れに属するかを、直感的に判断する機能のことを言う。

また、人が実際にパタン認識の機能を働かせるためには、さきに挙げた学習と記憶の過程と機能が必要であることも明らかである。

以上をまとめると、パタン認識の機能に必須の要素として、つぎの項目が挙げられる。

*明星大学情報学部電子情報学科

1. 外界の事象、あるいは観測パタン

2. 既知のカテゴリ概念

3. 直感的判断(狭義の認識)

4. 学習

5. 記憶

これらの項目のうち、第二の既知のカテゴリ概念とは、普通は人のある程度の広がりを持ったグループ、あるいは社会で共有されている概念、すなわち記号化可能なものを指している。

このようなカテゴリ概念の例として、光学的文字認識、いわゆるＯＣＲの処理対象である各種の文字が挙げられる。この場合、文字は記号そのものであり、特定の文化を共有する社会や国家の中では、その概念も共有されている。パタン認識の対象としての文字は、素朴で単純な問題から、十分に複雑で困難な問題を構成することができ、格好の題材になっている。

現在、印刷や手書の漢字を読取るＯＣＲは、ある程度普及しており、限定した範囲であれば、実用性を有することが認められている。しかし、１９７０年頃、通産省の大型プロジェ

クト(いわゆるパタン大プロ)などで、印刷漢字の機械読み取りがテーマの一つに挙げられたとき、その実現可能性と技術的達成度について、危惧の念を抱かざるを得なかったのが率直なところである。

ここで述べるパタン認識の安定性に関する議論は、当時、相関を利用する印刷文字認識手法の研究に携わっていた筆者が、その延長として考察したものである。

2 背景

OCRにかぎらず、広くパタン認識の分野で、不変性あるいは安定性について議論された例は、きわめて少なく、観測パタンの正規化にオイラの変分原理を適用した飯島と、観測空間、あるいは特徴空間における座標変換にたいする不変な表現形式を議論した甘利の例があるだけである。

これらの研究は、きわめて精緻なもので、その論理的完結性については疑念を抱く余地はない。しかし、前者が導いたモード関数、およびモード関数展開(KL展開と等価とされる)は、現実の問題には全く通用しないし、後者の結論は、すべてのアフィン変換に不変な表現形式は、零変換しかないという、いわば無駄な努力を防止するものでしかなかった。

これらの研究は、１９６０年代半ばに行われたもので、理論の出発点となる仮定/前提も、当時としては無意味なものとは考えられないが、今後さらに有用なパタン認識の理論を構築するには、これらの議論とは異なった見地から、パタン認識の性質を考察する必要がある。

3 パタン認識結果の安定性

ここで、パタン認識系とは、パタン認識を行う人、あるいはその機能の一部を限定的に模擬する機械(例えばＯＣＲ)を指すものとする。

このパタン認識系について、つぎのような性質を仮定する。

1. 認識できるカテゴリ概念をC_iとし、$i: 1..n$、nは十分に大きいとする
2. 学習は、iについて逐次的に行われるものとする
3. 認識系が、未知パタンXをカテゴリC_iと認識することを、つぎのように表記する

$$\psi(X, C_1, C_2, \cdots, C_n) \Rightarrow c_i \tag{1}$$

ここでc_iは、カテゴリC_iを代表する記号とする。

すなわち、カテゴリ概念C_iと記号c_iは、一対一で対応する。

式1で、$C_1, C_2, \cdots, C_n$は、同時に学習済みのカテゴリ概念の全体を示している。また、c_iは、パタン認識系の一般的な性質から、式2も成り立つものとする。

$$c_i : 1 \leq i \leq n \tag{2}$$

4. パタン認識系の認識機能は、新たなパタン概念の学習について安定である。

すなわち、

$$\psi(X, C_1, C_2, \cdots, C_n) \Rightarrow c_i \tag{3}$$

であるとすると、新たなパタン概念C_{n+1}を学習することによる、未知パタンXに対するパタン認識系の認識結果は、つぎの何れかである。

$$\left.\begin{array}{l} \psi(X, C_1, C_2, \cdots, C_n, C_{n+1}) \ \Rightarrow c_i \\ \text{または} \\ \psi(X, C_1, C_2, \cdots, C_n, C_{n+1}) \ \Rightarrow c_{n+1} \end{array}\right\} \tag{4}$$

式4の性質を、パタン認識系の学習過程における安定性と呼び、パタン認識系が満たすべき、基本的な性質であるとする。

このような安定性が、人の通常のパタン認識の振る舞いの中で、ごく自然なものであることは、具体的な例を考えて見れば容易に頷けよう。

たとえば、文字を始めて学習するさいに、まず算用数字を覚え、ついで英字を学習する過程を踏むものとする。

このような場合、算用数字だけ学習した段階では、0(ゼロ)と認識していた未知パタンを,英字を併せて認識出来るように学習した結果、O(オー)やB(ビー)と認識するように変わることはあっても、同じ算用数字の中で、認識結果が8(ハチ)に変わるようなことは、実際問題として起らないと言うことを意味している。

もし、ここで述べたパタン認識系の学習過程における安定性が成立しないとすれば、少なくとも数千字以上の漢字を認識しなければならない、漢字文化圏の人々にとって、文字の習得は想像を絶する事業となることであろう。

4 対判定

任意の二個一対のカテゴリC_i, C_j(ただし$i \neq j$)について、未知パタンXに対する識別規則(どちらにより近いかを判断する)を、つぎのように書く。

$$\left.\begin{array}{lll}\Phi(X,C_i,C_j) & \Rightarrow & c_i \\ \text{または} & \Rightarrow & c_j\end{array}\right\} \tag{5}$$

$j \neq i$であるすべてのjについて、式6が成立するとする。

$$\Phi(X,C_i,C_j) \Rightarrow c_i \tag{6}$$

このとき、未知パタンXのカテゴリをc_iと認識することにすれば、この識別規則の組み合わせで認識するパタン認識系は、上記の学習過程における安定性を満たす。

この$\Phi(X,C_i,C_j)$を、対判定(規則)と呼ぶ。

さらに、s_iを実数とし、未知パタンXとカテゴリC_iとの類似性を、式7の形で評価できるとする。

$$\phi(X,C_i) = s_i \tag{7}$$

$$\max_{k=1}^{n} s_k = s_i \tag{8}$$

さらに式8が成立し、$j \neq i$のとき$s_i > s_j$ならば、未知パタンXのカテゴリをc_iと認識することにする。

このようなパタン認識系は、式5、式6の対判定方式の特別な場合と考えることができ、周知の類似度法など、ＯＣＲの商用機のほとんどが採用している認識手法が該当する。

5 ある種のニューラルネットワークの解について

パーセプトロンに代表される、バックプロパゲーション型のニューラルネットワークは、一種の最小二乗近似をシミュレートする。

このようなモデルは、一般には三層以上の多層モデルで、格別の機能を有すると主張されることがある。以下、このような場合について、先に述べた安定性の見地から検討してみる。

簡単のため、モデルは三層とし、第一層がパタンの入力、第三層が認識結果の出力を行うものとする。第二層と第三層を構成する素子の個数はおのおのR個、T個とする。

このモデルで、各カテゴリC_iに属するサンプルパタンを第一層に入力し、なんらかの学習規則にもとづて学習した結果、第三層の対応する素子、たとえば第i番目の素子が排他的にオン(論理値1)になるとする。

このとき、第二層上にカテゴリC_iに対応する平均的なパタン$p_i(r)$が形成されるとすれば、所望の結果を与える第二層と第三層の素子間の結合係数は、$p_i(r)$と$p_j(r)$間の相互相関量$S_{i,j}$を用いて、容易に書き下すことが出来る。

$$S_{i,j} = \sum_{r=1}^{R} p_i(r) \times p_j(r) \tag{9}$$

$$w_r^t = (\sum_{i=1}^{n} p_i(r) \times \Delta_{i,t})/\Delta \tag{10}$$

ただし、Δは$S_{i,j}$の相互相関行列式の値、$\Delta_{i,t}$は、おなじ相互相関行列式の第i,t要素の余行列式の値を意味する。

すなわち、

$$\Delta = \begin{vmatrix} S_{1,1} & S_{1,2} & \cdot & \cdot & S_{1,n} \\ S_{2,1} & S_{2,2} & \cdot & \cdot & S_{2,n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ S_{n,1} & S_{n,2} & \cdot & \cdot & S_{n,n} \end{vmatrix} \tag{11}$$

また、式10は形式的に、式12の形に書くことが出来る。

$$\Delta \times w_r^t = \begin{vmatrix} S_{1,1} & S_{1,2} & \cdot & S_{1,t-1} & p_1(r) & S_{1,t+1} & \cdot & S_{1,n} \\ S_{2,1} & S_{2,2} & \cdot & S_{2,t-1} & p_2(r) & S_{2,t+1} & \cdot & S_{2,n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ S_{n,1} & S_{n,2} & \cdot & S_{n,t-1} & p_n(r) & S_{n,t+1} & \cdot & S_{n,n} \end{vmatrix} \tag{12}$$

式10で求めた第二層と第三層間の結合係数w_r^tを用いると、カテゴリC_iに対応する第2層上の平均パタン$p_i(r)$に対する第三層の出力は、w_r^tの作り方から容易にわかるように、式13で与えられる。

$$\sum_{r=1}^{R} w_r^t \times p_i(r) = \delta_{i,t} \tag{13}$$

ただし、$\delta_{i,t}$は、$i=t$のとき$\delta_{i,t}=1$、$i \neq t$のとき、$\delta_{i,t}=0$とする。

この場合、その作り方から明らかなように、w_r^tを用いた識

別行為は、一般的には学習過程の安定性の要請を満たさないことになる。

6 結 論

以上、パタン認識における不変性、もしくは学習過程における認識結果の安定性の概念の意味について述べた。また、この尺度を使うと、バックプロパゲーション型のニューラルネットワークに、パタン認識システムとして望ましい機能を期待することは困難であることを示した。

ここで述べた議論が、きわめて粗いものであることは承知しているが、今後の現実的パタン認識理論構築のきっかけの一つともなれば幸いである。