

パターンモデルを用いた不動点探索形 連想記憶システム方程式

鈴木 昇 一, 佐久間 拓 也

An Associative Memory System Equation of Fixed-Point Searching Type Using a Pattern-Model

Shoichi Suzuki and Takuya Sakuma

Abstract

It is presented here that an associative neural-network binary memory system on n -dimensional Euclidean space R^n suggested by K. Nijima is able to be extended to a system on any separable Hilbert space \mathfrak{H} using a corresponding pattern model $T\varphi$ of a pattern φ proposed by S. Suzuki. Applying to a problem concerning at how many stages a pattern-recognition process converges practically an iterative scheme suggested by P. Alfeld which can give a solution of a fixed-point equation on R^n , it is shown that the process has a solution of a constrained minimization problem.

By virtue of this research, the associative system may act upon unitary-transformation invariances such as rotations, scalings, and translations, etc. or perceptual constancies appearing in the psychology, and a design of its weights and thresholds and a evaluation about a convergence of the associatively recalling process may be precisely shown.

要 約

可分な Hilbert 空間 \mathfrak{H} の元としてのパターン φ の, S. Suzuki の提案したパターンモデル $T\varphi$ を用いれば, ユークリッド空間 R^n での新島耕一による“連想記憶のニューラルネット 2 値モデル”が \mathfrak{H} へと拡張され, 然も, R^n から R^n への写像に関する Peter Alfeld の“不動点方程式の解法としての an iterative scheme”を適用して, 本連想記憶モデルによるパターン認識過程が何段階でほぼ収束するかなどに関し, the constrained minimization problem の解として得られることが示されている。

本研究によって, 可分な Hilbert 空間 \mathfrak{H} での連想形記憶ニューラルネットの 2 値モデルにユニタリ座標変換不変性 (rotation, scaling, translation 等に対する不変性, あるいは心理学でいう一

種の知覚の恒常性)を備えさせることが可能になり、その重み・閾値の設計法、連想過程の収束の評価が精密にできるようになった。

1. まえがき

高々可算個の入出力例のなす集合(訓練例の集合)を用いて、希望する入出力関係を学習して、未知の入力に対しても“ある意味で最適な出力”をもたらしたり (associative memory models), あるエネルギー関数に極小値を与える入力変数値の組合せを求めたり⁽²⁸⁾ (neural network model for combinatorial optimization) するような

deterministic Hopfield network⁽²⁸⁾, backpropagation-learning network⁽²⁹⁾, recurrent network⁽³⁰⁾, stochastic Hopfield network としての Boltzmann machine などのニューラルネットによる情報処理とは^{(17), (20), (23)}, 結局は、入力から, “spurious attractors ではなくして、入力に最も関連している memorized attractors as fixed points” を呼び起こすことであり,

local minima ではなくして one global minimum を与える離散値をとる多変数の値の組合せを求めることである。

本研究の目的は、K. Nijijima による an associative model としての不動点収束形ニューラルネット⁽¹⁾の動作についての証明を完全な形で与え (2 付録A, B), このニューラルネットのユークリッド空間での動作 (有限次元動作) を、S. Suzuki の提案したモデル構成作用素 (model-construction operator)^{(3), (7), (8)~(11), (13), (16)~(18), (20), (21)} $T: \Phi \rightarrow \Phi$ を用い (付録C), ヒルベルト空間 \mathfrak{H} の (ある場合には、ユニタリ座標変換の下で不変な) (無限次元動作) に転換し、然も、P. Alfeld による不動点方程式の解法としての an iterative scheme for solving the constrained minimization problem⁽²⁾を適用して、その収束を評価することである。

2. 連想形ニューラルネットのシステム方程式

2. 1 ニューラルネットの記憶内容

可分な Hilbert 空間 \mathfrak{H} での内積, ノルムを各々, $(\cdot, \cdot), \|\cdot\| \equiv \sqrt{(\cdot, \cdot)}$ としよう。例えば、内積 (φ, η) が

$$(\varphi, \eta) = \int_M dm(x) \varphi(x) \cdot \bar{\eta}(x)$$

ここに、 $\bar{\eta}$ は η の複素共役であり、

$M: n$ 次元ユークリッド空間 R^n の可測部分集合

$dm(x)$: 正值 Lebesgue-Stieltjes 式測度 (1)

と与えられる可分な Hilbert 空間 $\mathfrak{H} = L_2(M; dm)$ を用意すればよい。

$$(\phi_k, \phi_\ell) = 0 \text{ if } k \neq \ell, = \|\phi_k\|^2 > 0 \text{ if } k = \ell \text{ (orthogonality)} \quad (2)$$

を満たす直交系 $\{\phi_k\}_{k \in L}$ と、特徴抽出写像

$u: \Phi \times L \rightarrow R^+$ (非負実数全体の集合)

ここに、 Φ は処理対象とするパターン $\varphi \in \mathfrak{H}$ の集合 (零元を含むような \mathfrak{H} の部分集合) であり (文献(20)の第24部を参照), $u(\varphi, k) \in R^+$ はパターン $\varphi \in \Phi$ から抽出され

る第 $k \in L$ 番目の特徴量 (3)

とを用いて, パターン $\varphi \in \Phi$ の代替物 (φ の正規化パターン)

$$T\varphi \equiv \sum_{k \in L} u(\varphi, k) \cdot \phi_k \in \Phi \quad (4)$$

が導入される (3), (7), (8) ~ (11), (13), (16) ~ (18), (20), (21)。

n 個のニューロンから成るニューラルネットを想定し, そのシステム方程式を

$$S_i \eta = \sum_{k \in L} g(\sum_{j=1}^n W_{ij}(k) \cdot u(\eta_j, k) - h_i(k)) \cdot \phi_k, \quad i = 1 \sim n \quad (5)$$

ここに, t は転置の意として,

$$\underline{\eta} = {}^t(\eta_1 \eta_2 \cdots \eta_n) \in \Phi^n \quad (6)$$

としよう。登場した諸記号の意味は次の通りである:

$$g(u) = [1 + \exp(-C \cdot u)]^{-1}, \quad c > 0$$

(sigmoidal function or logistic activation function) (ニューロン発火関数) (7)

$W_{ij}(k)$: 第 $k \in L$ 番目の特徴軸に関するニューロン j からニューロン i へのシナプス結合の重み

$h_i(k)$: 第 $k \in L$ 番目の特徴軸に関するニューロン i の閾値。

η_i : ニューロン i に記憶されているパターン

本ニューラルネットは, 写像 $S_i: \Phi^n \times W_i' \rightarrow \Phi$ の組

$$\underline{S} = (S_1 S_2 \cdots S_n): \Phi^n \times \prod_{i=1}^n W_i' \rightarrow \Phi^n \quad (8)$$

と同一視されることになる。ここに,

$$W = \{W_{ij}(k) \mid k \in L, i, j = 1 \sim n\}, \quad W_i = \{W_{ij}(k) \mid k \in L, j = 1 \sim n\}$$

$$h = \{h_i(k) \mid k \in L, i = 1 \sim n\}, \quad h_i = \{h_i(k) \mid k \in L\}$$

$$W' = W \cup \{h\}, \quad W_i' = W_i \cup \{h_i\}$$

Hilbert 空間 \mathfrak{H} の部分集合 Φ の n 個の直積 Φ^n の上で動作する以上の neural network に,

$\eta_i^k \in \Phi$: ニューロン i に記憶される第 $k (= 1 \sim m)$ 番目のパターン

として, 各

$$\underline{\eta}^k = {}^t(\eta_1^k \eta_2^k \cdots \eta_n^k) \in \Phi^n \quad (9)$$

を記憶させよう。それには, 不等式

$$0 < p < 2^{-1} < 1 - p < 1 \quad (10)$$

を満たす正実数 p を想定し, 不動点方程式の系

$$\begin{aligned} & \forall k (= 1 \sim m), \quad \forall i (= 1 \sim n), \quad \forall q \in L, \quad g(\sum_{j=1}^n W_{ij}(q) \cdot u(\eta_j^k, q) - h_i(q)) \\ & = u(\eta_i^k, q) \in \{p, 1 - p\} \end{aligned} \quad (11)$$

が満たされるように,

重み $W_{ij}(q)$
 閾値 $h_i(q)$, $q \in L, i, j = 1 \sim n$

を決定すればよい。

式(11)が成立すれば,

$$S_i \underline{\eta}^k = T \eta_i^k \in \Phi, i = 1 \sim n, k = 1 \sim m \quad (12)$$

がいえ,

$$\underline{S} \underline{\eta}^k = T \underline{\eta}^k \quad (13)$$

が成立することになるからである。ここに,

$$T \underline{\eta}^k = (T \eta_1^k \ T \eta_2^k \ \cdots \ T \eta_n^k) \in \Phi^n \quad (14)$$

と約束している。もし,

$$T \eta_i^k = \eta_i^k \in \Phi, i = 1 \sim n, k = 1 \sim m$$

(fixed-point equations) (15)

であるような embedded pattern η_i^k を想定していれば, 式(13)は

$$\underline{S} \underline{\eta}^k = \underline{\eta}^k \in \Phi^n \text{ あるいは } \underline{S} T \underline{\eta}^k = T \underline{\eta}^k \in \Phi^n$$

(fixed-point equations) (16)

と書き換えられ, 式(11)を満たす, 式(8)でいう本ニューラルネット \underline{S} は
 写像 S_i の組 \underline{S} の不動点として, $T \underline{\eta}^k$ を記憶している
 という解釈が得られ, 都合が良い。

2. 2 ニューラルネットによる想起認識

式(8)でいうニューラルネット \underline{S} に, 入力

$$\varphi = {}^t(\varphi_1 \ \varphi_2 \ \cdots \ \varphi_n) \in \Phi^n \quad (17)$$

が加えられたとき, 想起過程 (associatively-recalling process)

$$\begin{aligned} \varphi &\rightarrow S \varphi \rightarrow S^2 \varphi \equiv S(S \varphi) \rightarrow \cdots \\ &\rightarrow \cdots \rightarrow S^t \varphi \equiv S(S^{t-1} \varphi) \end{aligned} \quad (18)$$

が時点 $t \in \{0, 1, \cdots\}$ で収束するとは,

$$\begin{aligned} \varphi \langle t \rangle |_{t=0} &\equiv \varphi \\ \varphi \langle t \rangle &\equiv S^t \varphi \end{aligned} \quad (19)$$

として, 不動点方程式

$$\underline{S}(\varphi \langle t \rangle) = \varphi \langle t \rangle \quad (20)$$

が成立することをいう。そして、式(15)を満たす embedded pattern η_i^k について、

$$\forall i \in \{1, 2, \dots, n\}, \exists k \in \{1, 2, \dots, m\}, \varphi_i \langle t \rangle = T\eta_i^k \quad (21)$$

が成立するならば、

入力 $\varphi \in \Phi^n$ 内の第 i 成分パターン $\varphi_i \in \Phi$ は第 $k \in \{1, 2, \dots, m\}$ 番目のカテゴリ (category; 類概念) \mathcal{C}_k に認識推断される

という。

このように、記憶内容 $\eta^k \in \Phi^n$ 内の各 $\eta_i^k \in \Phi (i=1 \sim n)$ は

第 $k (= 1 \sim m)$ 番目のカテゴリ \mathcal{C}_k の第 i 番目の代表パターンである

と解釈されることになり、各カテゴリ \mathcal{C}_k に複数個 (n 個) の代表パターン η_i^k を用意することが可能になる。よって、例えば、視点の位置によって多様な形状を呈する物体の、各々の視点からみた代表パターンを用意できるなどの利点がこれ迄のニューラルネットとは異なり、従うことになる。また、入力 $\varphi \in \Phi^n$ 内の各 φ_i を同一のパターン φ と設定することにより ($\varphi_i = \varphi, i=1 \sim n$)、入力パターン $\varphi \in \Phi$ の多様な変形に応じた認識処理が可能となる。いかにすれば、例えば、aspect (物体の見え方) 認識に適用すれば、

入力パターン $\varphi \in \Phi$ は各 $\eta_i^k (i=1 \sim n)$ のいずれか (例えば、 η_j^k) と似ており、どの $\eta_j^\ell (\ell \neq k, j=1 \sim n)$ とも似ていなければ、 $T\eta_j^k$ として再生され、第 $k (= 1 \sim m)$ 番目のカテゴリ \mathcal{C}_k に認識推断されることになる。

3. 不動点方程式(11)の解

ニューロンの発火関数としての、式(7)の sigmoidal function $g(u)$ はニューラルネットの transfer function として採用されているものであるが、その逆関数 $g^{-1}(u)$ は次のように表わされる：

$$g(u) = v \in (0, 1) \equiv \{y | 0 < y < 1\} \Leftrightarrow u = g^{-1}(v) = - (1/c) \cdot \log_e [(1-v)/v] \quad \square$$

よって、式(11)は

$$\forall k (= 1 \sim m), \forall i (= 1 \sim n), \forall q \in L, \sum_{j=1}^n W_{ij}(q) \cdot u(\eta_j^k, q) - h_i(q) = g^{-1}(u(\eta_i^k, q)) \quad (22)$$

ここに、

$$\begin{aligned} & g^{-1}(u(\eta_i^k, q)) \\ &= - (1/c) \cdot \log_e \frac{1 - u(\eta_i^k, q)}{u(\eta_i^k, q)} \\ &= \operatorname{sgn}(u(\eta_i^k, q) - 2^{-1}) \cdot \frac{1}{c} \log_e \frac{1 - u(\eta_i^k, q)}{u(\eta_i^k, q)} \\ &= \operatorname{sgn}(u(\eta_i^k, q) - 2^{-1}) \cdot \frac{1}{c} \log_e \frac{1-p}{p} \end{aligned} \quad (23)$$

であり、

$$\operatorname{sgn}(u) = +1 \text{ if } u > 0, = 0 \text{ if } u = 0,$$

$$= -1 \text{ if } u > 0$$

と書ける。式(23)の計算には、

$$u(\eta_i^k, q) \in \{p, 1-p\} \text{ for any } q \in L$$

であることを用いている。同様に、式(22)は、

$$\begin{aligned} & \forall k (= 1 \sim m), \forall i (= 1 \sim n), \forall q \in L, \sum_{j=1}^n V_{ij}(q) \cdot u(\eta_j^k, q) - t_i(q) \\ & = \text{sgn}(u(\eta_i^k, q) - 2^{-1}) \end{aligned} \quad (24)$$

と変形される。ここに、 $V_{ij}(q), t_i(q)$ は

$$\begin{aligned} W_{ij}(q) &= V_{ij}(q) \cdot \frac{1}{c} \cdot \log_e \frac{1-p}{p} \\ h_i(q) &= t_i(q) \cdot \frac{1}{c} \cdot \log_e \frac{1-p}{p} \end{aligned} \quad (25)$$

とおかれている。

次の定理3. 1は、 $\ell \in L$ を1つ任意に選び固定したとき、不動点方程式の系である(11)、つまり、式(22)あるいは式(24)を満たす重み $W_{ij}(\ell), h_i(\ell), j=1 \sim n$ を、memorized patterns である式(9)の $\underline{\eta}^k \in \Phi^n (k=1 \sim m)$ を用い、1つの $i (= 1 \sim n)$ を固定する度に決定する手法を示すものである。

[定理3. 1] (重み $W_{ij}(\ell)$, 閾値 $h_i(\ell)$ の決定定理)

添字 $\ell \in L$ を任意に1つ選び、固定する。

$u(\eta_j^k, \ell)$ を第 $j (= 1 \sim n)$ 成分とするベクトル $\underline{u}(\eta^k_\ell)$ の $k (= 1 \sim m)$ に関する集合

$$\underline{u}(\underline{\eta}^k_\ell) \equiv \{u(\eta_j^k, \ell) \mid j=1 \sim n\}, k=1 \sim m$$

は1次独立とする。

$$\sum_{i=1}^n u(\eta_i^a, \ell) \cdot u(\eta_i^b, \ell)$$

を第 a 行第 b 列の要素とする行列を A_{11} として、
 m 次元横ベクトル

$$A_{21} = (2 \ 2 \ \cdots \ 2)$$

と、その転置ベクトル $A_{12} = {}^t A_{21}$ とを導入する。

$i (= 1 \sim n)$ を1つ任意に選び、固定する。

連立1次方程式

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{12} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \underline{s}_i \\ t_i \end{pmatrix} = \begin{pmatrix} \underline{s}_i \\ 0 \end{pmatrix}$$

ここに、

$$\underline{s}_i = {}^t (s_{i1} \ s_{i2} \ \cdots \ s_{im})$$

$$s_{ia} \equiv -2 \cdot \text{sgn}(u(\eta_i^a, \ell) - 2^{-1})$$

$$\underline{\xi} = {}^t(\xi_1 \ \xi_2 \ \cdots \ \xi_m)$$

は一意的な解 $\underline{\xi}, t_i$ を持ち、式(24)の $t_i(\ell)$ は

$$t_i(\ell) = t_i \quad (i = 1 \sim n)$$

と与えられる。また、等式

$$\sum_{j=1}^n v_{ij}(\ell) \cdot u(\eta_j^a, \ell) - t_i(\ell)$$

$$= \text{sgn}(u(\eta_i^a, \ell) - 2^{-1}), \quad a = 1 \sim m$$

の下で、

$$\sum_{j=1}^n v_{ij}(\ell)^2$$

を最小にする“最適化問題の解”

$$\underline{V}_i(\ell) = {}^t(v_{i1}(\ell) \ v_{i2}(\ell) \ \cdots \ v_{in}(\ell))$$

は、

$$v_{ij}(\ell)$$

$$= -2^{-1} \cdot \sum_{b=1}^m \xi_b \cdot u(\eta_j^b, \ell)$$

によって与えられる。よって、式(25)から、重み $W_{ij}(\ell)$ 、閾値 $h_i(\ell)$ が求まる。

(証明) 付録Aの定理A. 1を適用したものである。 □

4. 想起認識の収束

式(5)での、第 j ニューロンに記憶されているパターン η_j の第 $q \in L$ 番目の特徴量 $u(\eta_j, q)$ に注目し得られた量

$$u_q(\underline{\eta}) \equiv \{u(\eta_j, q) \mid j = 1 \sim n\}$$

での添字 $q \in L$ を1つ任意に固定して、付録Bでの式 (B. 2) の代りに、

$$F_i(\underline{\eta}, \varepsilon) \equiv F_i(\underline{\eta}, \varepsilon)(q)$$

$$\equiv g(\sum_{j=1}^n W_{ij}(q) \cdot u(\eta_j, q) - h_i(q))$$

$$+ \varepsilon \cdot v_i(q), \quad i = 1 \sim n$$

ここに、

$$v_i(q) = \frac{[\sum_{j=1}^n W_{ij}(q)^2 + h_i(q)^2] \cdot \|\phi_q\|^2}{\sum_{q \in L} [\sum_{j=1}^n W_{ij}(q)^2 + h_i(q)^2] \cdot \|\phi_q\|^2}$$

$$\underline{F}(\underline{\eta}, \varepsilon) = {}^t(F_1(\underline{\eta}, \varepsilon) \ F_2(\underline{\eta}, \varepsilon) \ \cdots \ F_n(\underline{\eta}, \varepsilon))$$

とおき、

$$u(\eta_j, q) \in \{p, 1-p\}$$

ここに, $0 < p < 2^{-1}$

と設定することにより, 付録 A の A, 2 節, 定理 A. 2 の系 (縮小写像定理) が適用可能となり, 従って, 付録 B での定理 B. 1 (COP の解定理) が適用可能である (付録 B の式 (B. 9) を参照)。

使用される諸記号が複雑になるだけで得られる諸性質は同様なものであるから, 書き直した具体的表現については割愛する。

5. むすび

新島耕一⁽¹⁾による“連想記憶モデルの収束”に関し精密かつ詳細な証明を与え (付録 A), Peter Alfeld⁽²⁾による“不動点問題を the constrained minimization problem へ転換する方法”について説明し (付録 B), S. Suzuki の提案した“ユニタリ座標変換不変性を備えたパターンモデル $T\varphi$ ”を介して (付録 C), 両研究者の成果を取り入れ可能な“(n 次元ユークリッド空間 R^n をその特別な場合として含む) 可分な一般抽象 Hilbert 空間 \mathcal{H} で動作する連想形記憶に関するニューラルネットの 2 値モデル”が提案された。新島, Peter なる両研究者の研究成果は共に, n 次元ユークリッド空間 R^n で得られている事態と比較すると, 可分な一般抽象 Hilbert 空間 \mathcal{H} での本研究成果の応用は広大となっていることに注意しておこう。

いいかえれば, 欠損部分のある入力パターンがその欠損が補なわれる形式で再生されるという新島の得た成果と, その再生過程の収束を制御できるという Alfeld の成果とを兼ね備えていることは勿論であるが,

- (イ) 有限次元ユークリッド空間 R^n での動作ではなしに, 関数空間 \mathcal{H} での動作が可能であり, 手を加えない形式での“ナマのパターン”をそのまま取り扱え, しかも
 - (ロ) 特別な細工を施すことなく, ユニタリ座標変換不変性 (心理学でいう一種の知覚の恒常性^{(10),(21)}) を備えた形式でのニューラルネットが構築可能である
- などが示されたことが本研究の成果であろう。

文 献

- (1) 新島耕一: “ある連想記憶モデルのダイナミクスとそのパターン認識への応用”, 電子情報通信学会技術研究報告 [ニューロコンピューティング], Vol. 90, No. 483, NC90-87, pp. 113-118 (1991-03)
- (2) Peter Alfeld: “Fixed Point Iteration with Inexact Function Values”, Mathematics of Computation, vol. 38, No. 157, pp. 87-98 (1982-01)
- (3) 鈴木昇一: “測度的不変量検出形認識系の構成理論”, 電子通信学会論文誌 (D), Vol. 55-D, No. 8, pp. 531-538 (1972-08)
- (4) 鈴木昇一: “手書き漢字の側抑制効果の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 (1974-12)
- (5) 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 (1975-04)
- (6) 鈴木昇一: “特徴量としての測度的ユニタリ不変量の完全な集合の一構成”, 電子通信学会論文誌 (D), Vol. J59-D, No. 9, pp. 678-680 (1976-09)
- (7) 鈴木昇一: “構造化情報パターンの 4 性質”, 電子通信学会論文誌 (D), Vol. J59-D, No. 12, pp. 937-938 (1976-12)

- (8) 鈴木昇一：“パターン認識における構造化モデルの4性質とその応用”，電子通信学会論文誌 (D)，Vol. J60—D，No. 9，pp. 710—717 (1977—09)
- (9) 鈴木昇一：“規格化特徴量の集合の完結構造モデルによる一意的決定”，電子通信学会論文誌 (D)，Vol. J60—D，No. 10，pp. 898—899 (1977—10)
- (10) 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 (1977—11)
- (11) 鈴木昇一：“構造モデル化写像の一般化”，電子通信学会論文誌 (A)，Vol. J66—A，No. 2，pp. 162—163 (1983—02)
- (12) 中村三郎，田代達也，鈴木昇一：“ソフトウェアをコンピュータに作らせる夢—1つの提案「MIS」について—”，コンピュータアクセス，pp. 54—62 (1990—01)
- (13) 鈴木昇一，斎藤静昭，奥野治雄，太田芳雄：“画像の復元とその計算機シミュレーション”，工学院大学研究報告，No. 39，pp. 198—206 (1976—01)
- (14) 鈴木昇一，太田芳雄，斎藤静昭，奥野治雄：“感覚空間回路の設計と作用素に対するラプラス変換法”，工学院大学研究報告，No. 40，pp. 122—134 (1976—06)
- (15) 鈴木昇一，柴山秀雄，福永一保，大本修，古田晋吾：“作用素に対するフーリエ変換法による側抑制特性の設計”，芝浦工業大学研究報告理工系編，Vol. 24，No. 1，pp. 147—155 (1980—03)
- (16) 鈴木昇一：“回転群と画像の分解・強調・構造化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)，Vol. 4，pp. 36—56 (1983—12)
- (17) 鈴木昇一：“連想形記憶器 MEMOTRON と日本語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)，Vol. 7，pp. 14—29 (1986—12)
- (18) 鈴木昇一：“認識プログラム FERT のリスト論的形式体系における表現”，情報研究 (文教大学情報学部)，Vol. 8，pp. 1—12 (1987—12)
- (19) 鈴木昇一：“収縮写像の構成用空間回路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)，Vol. 9，pp. 17—28 (1988—12)
- (20) 鈴木昇一：“パターン認識の数学的理論”，
 第Ⅰ部 (考え方，PRL 84-6，pp. 1—10，1984—05)，
 第Ⅱ部 (認識抽象と公理系，定理系，PRL 84-30，pp. 65—74，1984—09)，
 第Ⅲ部 (認識抽象と不動点諸定理，PRL 84-38，pp. 65—73，1984—09)，
 第Ⅳ部 (パターンの素領域，PRU 86-27，pp. 1—10，1985—09)，
 第Ⅴ部 (認識停止と認識同値，PRU 86-8，pp. 65—74，1986—05)，
 第Ⅵ部 (類似度関数の三構成法，PRU 86-35，pp. 51—60，1986—07)，
 第Ⅶ部 (類似度関数の実現と解析，PRU 87-69，pp. 1—8，1986—12)，
 第Ⅷ部 (大分類関数の自己組織化，PRU 87-1，pp. 1—8，1987—05)，
 第Ⅸ部 (帰属関数あいまい度と認識情報量，PRU 87-28，pp. 1—10，1987—07)，
 第Ⅹ部 (mixture 条件の研究，PRU 88-30，pp. 1—8，1988—07)，
 第Ⅺ部 (認識プログラム FERT_D の近似の鎖，PRU 89-1，pp. 1—8，1989—05)，
 第Ⅻ部 (ポテンシャル関数による認識過程の評価，PRU 89-27，pp. 1—8，1989—07)，
 第Ⅼ部 (認識プログラム FERT_D の不動点認識定理，PRU 89-40，pp. 1—8，1989—09)，
 第Ⅽ部 (線形帰属係数法と諸基本定理，PRU 89-66，pp. 1—8，1989—11)，
 第Ⅾ部 (パターンの構造的類似性をもたらず4種類の収縮写像，PRU 89-77，pp. 1—8，1989—12)，

- 第XVI部 (コネクショニスト・モデルと収縮写像, PRU89-136, pp. 9-16, 1990-03),
- 第XVII部 (ホップフィールドネットワーク 2 値モデルと収縮写像(1), PRU90-5, pp. 1-8, 1990-05),
- 第XVIII部 (ホップフィールドネットワーク 2 値モデルと収縮写像(2), PRU90-15, pp. 1-8, 1990-06),
- 第XIX部 (ホップフィールドネットワークの連続モデルと 2 種類の収縮写像(1), PRU90-29, pp. 9-16, 1990-07),
- 第XX部 (ホップフィールドネットワークの連続モデルと 2 種類の収縮写像(2), PRU90-125, pp. 1-8, 1991-02),
- 第XXI部 (誤差逆伝播ニューラルネットモデルと特徴抽出(1), PRU91-1, pp. 1-8, 1991-05),
- 第XXII部 (誤差逆伝播ニューラルネットモデルと特徴抽出(2), PRU91-29, pp. 23-28, 1991-06),
- 第XXIII部 (誤差逆伝播ニューラルネットモデルと特徴抽出(3), PRU91-42, pp. 1-8, 1991-07),
- 第XXIV部 (再帰領域方程式と標本化, PRU92-1, pp. 1-8, 1992-05),
- 第25部 (画像前処理, PRU92-18, pp. 1-8, 1992-06),
- 第26部 (線形歪を持った多次元パターンの, モーメントによる正規化, PRU92-25, pp. 1-8, 1992-09)
- 第27部 (モデル構成作用素による, *Extended Dynamic Axes Warping*(1), PRU92-89, pp. 1-8, 1992-12),
- 第28部 (モデル構成作用素による, *Extended Dynamic Axes Warping*(2), PRU92-102, pp. 1-8, 1993-01),
- 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]

(1984-05~1993-01)

- (21) 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. 11, pp. 2220-2238 (1994-11)
- (22) スミルノフ: “高等数学教程 (V 巻第二分冊)”, 小松彦三郎訳, pp. 256-263 (86, 87節), 共立出版 (1966-03)
- (23) Shun-ichi Amari: “Mathematical Foundations of Neurocomputing”, Proc. IEEE, Vol. 78, No. 9, pp. 1443-1463 (1990-09)
- (24) Naomi Blatt and Jacob Rubinstein: “The Canonical Coordinates Method for Pattern Recognition - II. Isomorphisms with Affine Transformations”, Pattern Recognition, Vol. 27, No. 1, pp. 99-107 (1994)
- (25) Gregory Gheen: “Distortion Invariant Volterra Filters”, Pattern Recognition, Vol. 27, No. 4, pp. 569-576 (1994)
- (26) Cem Yüceer and Kemal Oflazer: “A Rotation, Scaling, and Translation Invariant Pattern Classification System”, Pattern Recognition, Vol. 26, No. 5, pp. 687-710 (1993)
- (27) スミルノフ: “高等数学教程 (I 巻第 2 分冊)”, 山崎三郎訳, pp. 381-388 (167-169節), 共立出版 (1966-09)
- (28) Michael Fleisher: “The Hopfield Model with Multi-level Neurons”, Neural Information Processing Systems (Dever, CO, American Institute of Physics), pp. 278-289 (1987)
- (29) David E. Rumelhart, Geoffrey E. Hinton: “Learning Representation by back-propagating errors”, Nature, Vol. 323-9, pp. 533-536 (1986)
- (30) Masa-aki: “A Learning Algorithm to Teach Spatiotemporal Patterns to Recurrent Neural Networks”, Biological Cybernetics, Vol. 62, No. 3, pp. 259-263 (1990)
- (31) David H. Ackley, Geoffrey E. Hinton and Terrence J. Sejnowski: “A Learning Algorithm for Boltzmann Machines”, Cognitive Science, Vol. 9, pp. 147-169 (1985)
- (32) 鈴木昇一: “認識工学 (上)”, 柏書房 (1975-02)

付録A (新島⁽¹⁾の、 R^n でのニューラルネット連想記憶システムに関する2値モデルの挙動に関する証明)

新島⁽¹⁾の証明、解析は簡単過ぎる。少し整理した形式で詳細に証明、解析をやり直しておく。

A. 1 重み、閾値の決定

t は転置の意として、

$$\underline{x} = {}^t(x_1 \ x_2 \ \cdots \ x_n), \ \underline{z} = {}^t(z_1 \ z_2 \ \cdots \ z_n) \in R^n \text{ (} n \text{次元ユークリッド空間)}$$

に対し、

$$\text{内積 } (\underline{x}, \underline{z}) = \sum_{i=1}^n x_i \cdot z_i$$

$$\text{ノルム } \|\underline{x}\| = \sqrt{(\underline{x}, \underline{x})}$$

を導入する。

まず、次の定理A. 1を証明する。

[定理A. 1] (重み v_{ij} , 閾値 t_i の決定定理)

各 x_i^a は実数値として、実数値ベクトル \underline{x}^a の集合

$$\underline{x}^a = {}^t(x_1^a \ x_2^a \ \cdots \ x_n^a), \ a = 1, 2, \dots, m$$

は1次独立な系とする。内積

$$(\underline{x}^a, \underline{x}^b) = \sum_{i=1}^n x_i^a \cdot x_i^b$$

を第a行第b列の要素とする行列を A_{11} とし、 m 次元横ベクトル $A_{21} = (2 \ 2 \ \cdots \ 2)$ と、その転置ベクトル $A_{12} = {}^t A_{21}$ とを導入する。

$i(=1 \sim n)$ を任意に1つ選び、固定する。

連立1次方程式

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & 0 \end{pmatrix} \begin{pmatrix} \underline{\xi} \\ t_i \end{pmatrix} = \begin{pmatrix} \underline{s}_i \\ 0 \end{pmatrix} \tag{A. 1}$$

ここに、

$$\underline{s}_i = {}^t(s_{i1} \ s_{i2} \ \cdots \ s_{im})$$

$$s_{ia} \equiv -2 \cdot \text{sgn}(x_i^a - 2^{-1})$$

$$\text{sgn}(u) = +1 \text{ if } u > 0, = 0 \text{ if } u = 0$$

$$= -1 \text{ if } u < 0.$$

$$\underline{\xi} = {}^t(\xi_1 \ \xi_2 \ \cdots \ \xi_n)$$

は、一意的な解 $\underline{\xi}, t_i$ をもち、等式

$$(\underline{x}^a, V_i) - t_i = \text{sgn}(x_i^a - 2^{-1}), \ a = 1, 2, \dots, m \tag{A. 2}$$

の下で、

$$\| \underline{V}_i \|^2 (= (V_i, V_i) = \sum_{j=1}^n v_{ij}^2) \quad (\text{A. 3})$$

を最小にする最適化問題の解

$$V_i = {}^t(v_{i1} \ v_{i2} \ \cdots \ v_{in})$$

は

$$v_{ij} = -2^{-1} \sum_{b=1}^m \xi_b \cdot x_i^b \quad (\text{A. 4})$$

によって与えられる。 □

上述の定理 A. 1 を証明する前に、先ず、ラグランジュの乗数法について説明しておこう。

ラグランジュの乗数法について ⁽²⁷⁾

$(m+n)$ 個の変数 $x_i, i=1 \sim m+n$ の間に、

$$\varphi_i(x_1, x_2, \dots, x_m, x_{m+1}, \dots, x_{m+n}) = 0, i=1 \sim n \quad (\text{A. 5})$$

という関係があるとき、関数

$$f(x_1, x_2, \dots, x_m, x_{m+1}, \dots, x_{m+n}) \quad (\text{A. 6})$$

の (条件付き) 極大・極小を求めることを考えよう。

今、式 (A. 5) を n 個の変数、例えば、

$$x_{m+1}, x_{m+2}, \dots, x_{m+n} \quad (\text{従属変数の値})$$

について解けば、これらは残りの変数

$$x_1, x_2, \dots, x_m \quad (\text{独立変数の組})$$

で表される。これらを、式 (A. 6) の関数 f に代入すれば、 m 個の独立変数 x_1, x_2, \dots, x_m の関数ができ、よって、条件付極大・極小の問題は、条件なしの極大・極小を求めることに帰着する。

上述の条件付極大・極小問題の解を求めるには、新たに、関数

$$\Phi \equiv f + \sum_{k=1}^n \lambda_k \cdot \varphi_k \quad (\text{A. 7})$$

において、 $\lambda_1, \lambda_2, \dots, \lambda_n$ を定数 (ラグランジュの乗数) と考え、これをすべての $x_s (s=1, 2, \dots, m, m+1, \dots, m+n)$ について偏微分したものを 0 に等しいとおけばよい。

いいかえれば、

$$\begin{aligned} \partial \Phi / \partial x_s &= \partial f / \partial x_s + \sum_{i=1}^n \lambda_i \cdot (\partial \varphi_i / \partial x_s) \\ &= 0, s=1, 2, \dots, m \end{aligned} \quad (\text{A. 8})$$

$$\begin{aligned} \partial \Phi / \partial x_s &= \partial f / \partial x_s + \sum_{i=1}^n \lambda_i \cdot (\partial \varphi_i / \partial x_s) \\ &= 0, s=m+1, m+2, \dots, m+n \end{aligned} \quad (\text{A. 9})$$

を求めれば、後半の式 (A. 9) が $\lambda_1, \lambda_2, \dots, \lambda_n$ を決定する方程式であり、2式 (A. 8), (A. 9) と

$$\partial\Phi/\partial\lambda_q = \varphi_q = 0, q = 1 \sim n \quad (\text{A. 10})$$

とを満たす変数の組 $x_1, x_2, \dots, x_m, x_{m+1}, \dots, x_{m+n}$ を求めればよい。

このラグランジュの乗数法の制限について指摘しておこう。

方程式 (A. 9) から, $\lambda_1, \lambda_2, \dots, \lambda_n$ を決定できると仮定していることに注意しなければならない。従って, このラグランジュの乗数法によると, 条件付極大・極小になる点 x_1, x_2, \dots, x_m の全部が出て来ないかも知れない。

(定理 A. 1 の証明)

Lagrange の未定定数法 (上述) を適用しよう。

$\xi_b (b = 1 \sim m)$ を Lagrange の乗数として,

$$J \equiv \sum_{j=1}^n v_{ij}^2 + \sum_{b=1}^m \xi_b \cdot [\sum_{j=1}^n x_j^b v_{ij} - t_i - \text{sgn}(x_i^b - 2^{-1})]$$

とおく。 $\partial J/\partial \xi_a = 0, a = 1 \sim m$ より,

$$\sum_{j=1}^n x_j^a v_{ij} - t_i - \text{sgn}(x_i^a - 2^{-1}) = 0, a = 1 \sim m \quad (\text{A. 11})$$

が得られ, まな, $\partial J/\partial v_{ij} = 0, j = 1 \sim n$ より,

$$2v_{ij} + \sum_{b=1}^m \xi_b \cdot x_j^b = 0, j = 1 \sim n \quad (\text{A. 12})$$

が得られ, $\partial J/\partial t_i = 0$ より,

$$\sum_{b=1}^m \xi_b \cdot (-1) = 0 \quad \therefore \sum_{b=1}^m \xi_b = 0 \quad (\text{A. 13})$$

が得られる。式 (A. 12) を解いて, v_{ij} を求めると,

$$v_{ij} = -2^{-1} \sum_{b=1}^m \xi_b \cdot x_j^b \quad (\text{A. 14})$$

が得られ, この式 (A. 14) を式 (A. 11) に代入すると,

$$\begin{aligned} & \sum_{b=1}^m \xi_b \sum_{j=1}^n x_j^a x_j^b + 2t_i \\ & = -2 \cdot \text{sgn}(x_i^a - 2^{-1}), a = 1 \sim m \end{aligned} \quad (\text{A. 15})$$

が得られる。この式 (A. 15) と式 (A. 13) とを組合わせると, 連立 1 次方程式 (A. 1) が導かれる。

連立 1 次方程式 (A. 1) が一意的な解をもつことを示そう。

$\underline{x}^a, a = 1 \sim m$ は 1 次独立と仮定しているから, Gram の行列 A_{11} は正則であり, その行列式 $\det(A_{11})$ は

$$\det(A_{11}) \neq 0$$

である。また, 恒等式

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & 0 \end{pmatrix} \begin{pmatrix} E & -A_{11}^{-1}A_{12} \\ \underline{0} & 1 \end{pmatrix}$$

$$= \begin{pmatrix} A_{11} & \mathbf{0} \\ A_{21} & -A_{21}A_{11}^{-1}A_{12} \end{pmatrix}$$

ここに、 $\mathbf{0}$ は m 次元零ベクトル (横ベクトル) を表し、 E は m 次単位行列が成り立つ。従って、

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & \mathbf{0} \end{pmatrix}$$

とおくと、 A の行列式 $\det(A)$ に関し、

$$\det(A) = -\det(A_{11}) \cdot \det(A_{21}A_{11}^{-1}A_{12})$$

がいえる。 A_{11} は正定値行列であるから、 A_{11}^{-1} も正定値行列になるから、 A_{12} が横ベクトル A_{21} の転置ベクトルであることに注意すると、 $\det(A)$ の右辺は負になり、 0 にならない。

従って、連立1次方程式 (A. 1) は一意的な解をもつ。 □

A. 2 ニューラルネットの縮小写像性

さて、実数値 W_{ij}, h_i を

W_{ij} : ニューロン j からニューロン i へのシナプス結合の重み (synaptic weight)

h_i : ニューロン i の閾値

として、 n 個のニューロンから成すシステム方程式

$$y_i = g(\sum_{j=1}^n W_{ij}x_j - h_i), i = 1 \sim n$$

を考える。

$$\underline{x}^a = {}^t(x_1^a, x_2^a, \dots, x_n^a)$$

ただし、 t は転置の意

を第 a ($a = 1 \sim m$) 番目の不動点として、

$$x_i^a = g(\sum_{j=1}^n W_{ij}x_j^a - h_i), i = 1 \sim n \tag{A. 16}$$

が成立するように、各 W_{ij}, h_i をあらかじめ決定しておいて、

$$\underline{x}(t) |_{t=0} \equiv \underline{x} = {}^t(x_1, x_2, \dots, x_n)$$

から出発して、

$$x_i(t+1) = g(\sum_{j=1}^n W_{ij} \cdot x_j(t) - h_i), i = 1 \sim n, t \in \{0, 1, 2, \dots\}$$

を求めて行き、不動点方程式 (fixed-point equation)

$$x_i(t+1) = x_i(t), i = 1 \sim n$$

がある時刻 t で成立したとき、

\underline{x} から $\underline{x}(t) = {}^t(x_1(t), x_2(t), \dots, x_n(t))$ が想起されたとするニューラルネットを研究の対象としよう。

$$g(u) = 1/[1 + \exp[-cu]]$$

$$\in (0, 1) \equiv \{s | 0 < s < 1\}, \text{ここに } c > 0 \quad (\text{A. 17})$$

という一実変数 u の関数を導入すると,

$$g(u) = s$$

$$\Leftrightarrow u = g^{-1}(s) = -(1/c) \cdot \log_e[(1-s)/s] \quad (\text{A. 18})$$

がいえ, 不等式

$$0 < p < 2^{-1} < 1-p < 1 \quad (\text{A. 19})$$

を満たす2つの実数 $p, 1-p$ に対し

$$g^{-1}(p) = -(1/c) \cdot \log_e[(1-p)/p]$$

$$g^{-1}(1-p) = (1/c) \cdot \log_e[(1-p)/p] \quad (\text{A. 20})$$

が成り立つ。ここで, 変数変換

$$W_{ij} \equiv v_{ij} \cdot (1/c) \cdot \log_e[(1-p)/p]$$

$$h_i \equiv t_i \cdot (1/c) \cdot \log_e[(1-p)/p], i, j = 1 \sim n \quad (\text{A. 21})$$

を導入すると, 記憶ベクトル

$$\underline{x}^a = (x_1^a \ x_2^a \ \cdots \ x_n^a) \quad (\text{A. 22})$$

を不動点とするニューラルネットのシステム方程式

$$g(\sum_{j=1}^n W_{ij} \cdot x_j^a - h_i) = x_i^a \in \{p, 1-p\}$$

$$i = 1 \sim n, a = 1 \sim m \quad (\text{A. 23})$$

は,

$$\sum_{j=1}^n W_{ij} \cdot x_j^a - h_i = g^{-1}(x_i^a)$$

$$= \mp (1/c) \cdot \log_e[(1-p)/p]$$

$$= \text{sgn}(x_i^a - 2^{-1})$$

$$\cdot (1/c) \cdot \log_e[(1-p)/p]$$

$$i = 1 \sim n, a = 1 \sim m \quad (\text{A. 24})$$

と変形されるということから,

$$\sum_{j=1}^n V_{ij} \cdot x_j^a - t_i = \text{sgn}(x_i^a - 2^{-1})$$

$$i = 1 \sim n, a = 1 \sim m \quad (\text{A. 25})$$

ここに,

$$\text{sgn}(u) = +1 \text{ if } u > 0, = 0 \text{ if } u = 0, = -1 \text{ if } u < 0 \quad (\text{A. 26})$$

と書き直される。次の諸記法をも導入しておく。

$$\underline{x} = {}^t(x_1 \ x_2 \ \cdots \ x_n) \in R^n \text{ (} n \text{次元ユークリッド空間)} \quad (\text{A. 27})$$

$$G_i(\underline{x}) = g(\sum_{j=1}^n W_{ij} \cdot x_j - h_i) \quad (\text{A. 28})$$

$$G(\underline{x}) = {}^t(G_1(\underline{x}) \ G_2(\underline{x}) \ \cdots \ G_n(\underline{x})) \quad (\text{A. 29})$$

$$\begin{aligned} & \|G(\underline{x}) - G(\underline{z})\| \\ &= [\sum_{j=1}^n (G_j(\underline{x}) - G_j(\underline{z}))^2]^{1/2} \end{aligned} \quad (\text{A. 30})$$

$$V_i = {}^t(v_{i1} \ v_{i2} \ \cdots \ v_{in}) \quad (\text{A. 31})$$

$$\|V_i\| = [\sum_{j=1}^n V_{ij}^2]^{1/2} \quad (\text{A. 32})$$

□

このとき、本ニューラルネットの縮小写像性

「不等式 $0 < \kappa < 1$ を満たすある定数 κ が存在して、不等式

$$\|G(\underline{x}) - G(\underline{z})\| \leq \kappa \cdot \|\underline{x} - \underline{z}\|$$

が任意の $\underline{x}, \underline{z} \in R^n$ に対し成り立つ、」 (A. 33)

ような 2^{-1} より小さい正実数 p が存在することが、次の定理 A. 2 の系で示される。

[定理 A. 2] (ニューラルネットの出力ベクトル間の評価定理)

各 $v_{ij} (j=1 \sim n)$ は実数値として、

$$\begin{aligned} \sum_{j=1}^n v_{ij} \cdot x_j^a - t_i &= \text{sgn}(x_i^a - 2^{-2}), \ x_i^a \in \{p, 1-p\} \\ i &= 1 \sim n, \ a = 1 \sim m, \end{aligned} \quad (\text{A. 34})$$

を満たしているとする。記憶パターンの集合

$$\underline{x}^a = {}^t(x_1^a \ x_2^a \ \cdots \ x_n^a), \ a = 1 \sim m$$

内のいずれか 1 つの記憶パターンを \underline{x}^* とする。 \underline{x}^* を中心にもつ半径 δ の球

$$B_\delta(\underline{x}^*) = \{\underline{x} \in R^n \mid \|\underline{x} - \underline{x}^*\| \leq \delta\}$$

ここに、 $\|\underline{x} - \underline{x}^*\| = [\sum_{i=1}^n (x_i - x_i^*)^2]^{1/2}$ を用意する。このとき、不等式

$$0 < \rho < 1 \quad (\text{A. 35})$$

を満たす ρ に対し、

$$0 < \delta \leq \min_{i=1 \sim n} \rho / \|V_i\| = \rho / \max_{i=1 \sim n} \|V_i\| \quad (\text{A. 36})$$

よって、

任意の $i (= 1 \sim n)$ につき、

$$\|V_i\| \leq \min_{i=1 \sim n} \|V_i\| \leq \rho / \delta \quad (\text{A. 37})$$

が成り立つならば、任意の $\underline{x}, \underline{z} \in B_\delta(\underline{x}^*)$ に対し、不等式

$$\|G(x) - G(z)\| \leq \kappa \cdot \|x - z\| \quad (\text{A. 38})$$

が成り立つ。ここに、

$$\kappa = \{\sqrt{n}\rho/\delta\} \cdot p^{1-p} \log_e p^{-1} \quad (\text{A. 39})$$

[系] (縮小写像定理)

$$\lim_{p \rightarrow +0} \kappa = 0 \quad (\text{A. 40})$$

が成立し、 p が十分小さいと $\kappa < 1$ を得て、式 (A. 29) の写像 G は縮小写像 (contraction-mapping) となる。□

上記の定理 A. 2, その系を証明するために、5 命題 a. 1 ~ a. 5 を用意しよう。

[命題 a. 1] (平均値の定理)

関数 $y = f(x)$ が閉区間 $[a, b] \equiv \{x | a \leq x \leq b\}$ において連続であり、かつ 开区間 $(a, b) \equiv \{x | a < x < b\}$ である微分可能であるとき、等式

$$f(a+h) = f(a) + h \cdot (d/dx) f(x) |_{x=a+\theta h}$$

を満たす θ ($0 < \theta < 1$) が存在する。□

[命題 a. 2] (シグモイド関数の諸性質)

1 実変数 u の実数値関数

$$g(u) = [1 + \exp[-cu]]^{-1} \quad (c > 0)$$

について、次の(1)~(7)が成立する：

- (1) $g(0) = 2^{-1}$
- (2) $0 \leq g(u) \leq 1$
- (3) $\lim_{u \rightarrow -\infty} g(u) = 0, \lim_{u \rightarrow +\infty} g(u) = 1$
- (4) $(d/du)g(u) = c \cdot g(u) \cdot [1 - g(u)] \geq 0$
- (5) $\lim_{u \rightarrow \pm\infty} (d/du)g(u) = 0$
- (6) $s = g(u) \in \{y | 0 < y < 1\}$

を満たす $u = g^{-1}(s)$ は

$$u = g^{-1}(s) = - (1/c) \cdot \log_e [s^{-1}(1-s)]$$

(7) $0 \leq p, q \leq 1$ として、

$$\int_p^q ds g^{-1}(s)$$

$$= (1/c) \cdot [H(p) - H(q)]$$

ここに、

$$H(p) \equiv -p \log_e p - (1-p) \log_e (1-p)$$

$$H(q) \equiv -q \log_e q - (1-q) \log_e (1-q). \quad \square$$

[命題 a. 3]

上の命題 a. 2 の関数 $g(u)$ について,

$$g(u) \cdot [1 - g(u)] = \frac{1}{1 + \exp[-cu]} \cdot \frac{\exp[-cu]}{1 + \exp[-cu]}$$

□

[命題 a. 4]

$0 \leq p \leq 1 \wedge 0 \leq x \leq 1$ のとき, 不等式

$$[p + (1 - p)x]^{-1} \leq x^{-1}$$

が成り立ち, 等号は

$$p = 0 \vee x = 1$$

のときに限る。

□

[命題 a. 5]

関数 $h(x) = x(1 - x)$ for $0 \leq x \leq 1$ について,

$$(1) \max_{0 \leq x \leq 1} h(x) = 1/4 = h(2^{-1})$$

$$(2) \max_{0 \leq x \leq 1} h(x) = 0 = h(0) = h(1)$$

$$(3) 0 \leq x_1 \leq x_2 \leq 2^{-1} \text{ ならば,}$$

$$0 \leq h(x_1) \leq h(x_2) \leq 1/4$$

$$(4) 2^{-1} \leq x_1 \leq x_2 \leq 1 \text{ ならば,}$$

$$1/4 \geq h(x_1) \geq h(x_2) \geq 0.$$

□

(定理 A. 2 の証明)

$$G_i(x) - G_i(z) = g(\sum_{j=1}^n W_{ij} \cdot x_j - h_i) - g(\sum_{j=1}^n W_{ij} \cdot z_j - h_i)$$

であるから, $G_i(x)$ を $y_i \equiv \sum_{j=1}^n W_{ij} \cdot x_j - h_i$ の関数をみて, 命題 a. 1 で示される平均値の定理を適用すれば,

ある λ ($0 < \lambda < 1$) が存在して,

$$G_i(x) - G_i(z) = (d/du)g(u) \Big|_{u=y_i} \cdot [\sum_{j=1}^n W_{ij}(x_j - z_j)] \quad (\text{A. 41})$$

ここに,

$$y_i \equiv \sum_{j=1}^n W_{ij} \cdot [\lambda x_j + (1 - \lambda)z_j] - h_i \quad (\text{A. 42})$$

が成り立つ。式 (A. 42) の y_i は

$$y_i \equiv \sum_{j=1}^n W_{ij} \cdot x_j^* - h_i + \sum_{j=1}^n W_{ij} \cdot [\lambda(x_j - x_j^*) + (1 - \lambda)(z_j - x_j^*)] \quad (\text{A. 43})$$

と変形され、命題 a、2 の(4)を用いて、式 (A. 41) は

$$\begin{aligned} & G_i(\underline{x}) - G_i(\underline{z}) \\ &= c \cdot g(y_i) \cdot [1 - g(y_i)] \cdot [\sum_{j=1}^n W_{ij}(x_j - z_j)] \end{aligned} \quad (\text{A. 44})$$

と変形される。ここで、Schwarz の不等式

$$|\sum_{j=1}^n x_j \cdot \bar{y}_j| \leq [\sum_{j=1}^n |x_j|^2]^{1/2} \cdot [\sum_{j=1}^n |y_j|^2]^{1/2}$$

を適用すると、

$$\begin{aligned} & |G_i(\underline{x}) - G_i(\underline{z})| \\ &= c \cdot g(y_i) \cdot [1 - g(y_i)] \cdot |[\sum_{j=1}^n W_{ij}(x_j - z_j)]| \\ &\leq c \cdot g(y_i) \cdot [1 - g(y_i)] \cdot \|W_i\| \cdot \|\underline{x} - \underline{z}\| \\ &\quad \text{ここに、} W_i = {}^t(W_{i1} \ W_{i2} \ \cdots \ W_{in}) \end{aligned} \quad (\text{A. 45})$$

が成り立つ。

$\underline{x}, \underline{x}^* \in B_\delta(\underline{x}^*)$ であるから、

$$\begin{aligned} & \|\lambda(\underline{x} - \underline{x}^*) + (1 - \lambda)(\underline{z} - \underline{x}^*)\| \\ &\leq \lambda \cdot \|\underline{x} - \underline{x}^*\| + (1 - \lambda) \cdot \|\underline{z} - \underline{x}^*\| \\ &\leq \lambda\delta + (1 - \lambda)\delta = \delta \end{aligned}$$

を得て、これを Schwarz の不等式を適用して得られる不等式

$$\begin{aligned} & |W_i \cdot [\lambda(\underline{x} - \underline{x}^*) + (1 - \lambda)(\underline{z} - \underline{x}^*)]| \\ &\leq \|W_i\| \cdot \|\lambda(\underline{x} - \underline{x}^*) + (1 - \lambda)(\underline{z} - \underline{x}^*)\| \end{aligned}$$

に代入すれば、式 (A. 43) の右辺の部分式に関する不等式

$$\begin{aligned} & |\sum_{j=1}^n W_{ij} \cdot [\lambda(x_j - x_j^*) + (1 - \lambda)(z_j - x_j^*)]| \\ &\leq |W_i \cdot [\lambda(\underline{x} - \underline{x}^*) + (1 - \lambda)(\underline{z} - \underline{x}^*)]| \\ &\leq \|W_i\| \cdot \delta \end{aligned} \quad (\text{A. 46})$$

が得られる。

ここで、式 (A. 16) から

$$\sum_{j=1}^n W_{ij} \cdot x_j^* - h_i = g^{-1}(x_i^*), \quad i = 1 \sim n \quad (\text{A. 47})$$

が成立していることに注意して、

$$y_i^- \equiv \sum_{j=1}^n W_{ij} \cdot x_j^* - h_i - \delta \cdot \|W_i\| \quad (\text{A. 48})$$

$$y_i^+ \equiv \sum_{j=1}^n W_{ij} \cdot x_j^* - h_i + \delta \cdot \|W_i\| \quad (\text{A. 49})$$

とおくと、2式 (A. 43), (A. 46) より

$$y_i^- \leq y_i \leq y_i^+, \quad i = 1 \sim n \quad (\text{A. 50})$$

$$g(y_i^-) \leq g(y_i) \leq g(y_i^+), i = 1 \sim n \quad (\text{A. 51})$$

が成り立つことが知れた。ここに、

$$\delta \cdot \|W_i\| \leq \rho \cdot c^{-1} \log_e [p^{-1}(1-p)] \quad (\text{A. 52})$$

が成り立つことは、式 (A. 36)、式 (A. 21) を考慮すると、

$$\begin{aligned} & \delta \cdot \|W_i\| \\ & \leq [\rho / \max_{i=1 \sim n} \|V_i\|] \cdot \|W_i\| \\ & = [\rho / \max_{i=1 \sim n} \|V_i\|] \cdot c^{-1} \log_e [p^{-1}(1-p)] \cdot \|V_i\| \\ & \leq \rho \cdot c^{-1} \log_e [p^{-1}(1-p)] \end{aligned}$$

からわかる。

$$x_i^* = p \text{ or } 1-p$$

に従って、2つの場合 I、II に分けて、不等式

$$\begin{aligned} & g(y_i) \cdot (1 - g(y_i)) \\ & \leq p(1-p) \cdot \exp[c\delta \|W_i\|] \end{aligned} \quad (\text{A. 53})$$

が成り立つことを証明しよう。

I. $x_i^* = p$ のとき

式 (A. 49) の y_i^+ は、式 (A. 47) を代入して、

$$\begin{aligned} y_i^+ & = g^{-1}(x_i^*) + \delta \cdot \|W_i\| \\ & \leq -c^{-1} \log_e [p^{-1}(1-p)] + \rho \cdot c^{-1} \log_e [p^{-1}(1-p)] \\ & = [c^{-1} \cdot \log_e [p^{-1}(1-p)]] \cdot [\rho - 1] \\ & < 0 \quad \because \text{式 (A. 35)} \end{aligned}$$

を得て、よって、式 (A. 51) をも考慮して、

$$g(y_i) \leq g(y_i^+) < 2^{-1}$$

がいえる。従って、命題 a. 5 の(3)を適用して、

$$g(y_i) \cdot [1 - g(y_i)] \leq g(y_i^+) \cdot [1 - g(y_i^+)] \quad (\text{A. 54})$$

を得る。ここで、式 (A. 20) を考慮すると、

$$\begin{aligned} & \exp[-c^{-1} \cdot g^{-1}(x_i^*)] \\ & = \exp[\log_e [p^{-1}(1-p)]] \\ & = p^{-1}(1-p) \end{aligned} \quad (\text{A. 55})$$

であるから、

$$\begin{aligned} & \exp[-cy_i^+] \\ &= p^{-1}(1-p) \cdot \exp[-c\delta \|W_i\|] \end{aligned} \quad (\text{A. 56})$$

を得て、命題 a. 3 から、

$$\begin{aligned} & g(y_i^+) \cdot [1 - g(y_i^+)] \\ &= \frac{1}{1 + p^{-1} \cdot (1-p) \cdot \exp[-c \cdot \delta \cdot \|W_i\|]} \cdot \frac{p^{-1} \cdot (1-p) \cdot \exp[-c\delta \|W_i\|]}{1 + p^{-1} \cdot (1-p) \cdot \exp[-c \cdot \delta \cdot \|W_i\|]} \\ &= \frac{p}{p + (1-p) \cdot \exp[-c \cdot \delta \cdot \|W_i\|]} \cdot \frac{(1-p) \cdot \exp[-c \cdot \delta \cdot \|W_i\|]}{p + (1-p) \cdot \exp[-c \cdot \delta \cdot \|W_i\|]} \\ &= \frac{p(1-p)}{p + (1-p) \cdot \exp[-c\delta \|W_i\|]} \cdot \frac{1}{p \cdot \exp[c \cdot \delta \|W_i\|] + (1-p)} \\ &= \frac{p(1-p)}{p + (1-p) \cdot \exp[-c\delta \|W_i\|]} \cdot \frac{1}{1 + p[\exp[c \cdot \delta \cdot \|W_i\|] - 1]} \\ &\leq \frac{p(1-p)}{p + (1-p) \cdot \exp[-c\delta \|W_i\|]} \leq p(1-p) \cdot \exp[c \cdot \delta \|W_i\|] \\ &\quad \because \text{命題 a. 4} \end{aligned}$$

つまり、

$$g(y_i^+) \cdot [1 - g(y_i^+)] \leq p(1-p) \cdot \exp[c \cdot \delta \cdot \|W_i\|] \quad (\text{A. 57})$$

を得て、2式 (A. 54), (A. 57) より、式 (A. 53) の成立がわかる。

II. $x_i^* = 1-p$ のとき、

式 (A. 48) の y_i^- は、式 (A. 47) から

$$\begin{aligned} & y_i^- = g^{-1}(x_i^*) - \delta \cdot \|W_i\| \\ &= c^{-1} \log_e [p^{-1}(1-p)] - \delta \cdot \|W_i\| \quad \because \text{式 (A. 20)} \\ &\geq c^{-1} \log_e [p^{-1}(1-p)] - \rho \cdot c^{-1} \log_e [p^{-1}(1-p)] \quad \because \text{式 (A. 52)} \\ &= [c^{-1} \cdot \log_e [p^{-1}(1-p)]] \cdot [1 - \rho] \\ &> 0 \quad \because \text{式 (A. 35)} \\ &\therefore g(y_i^-) > 2^{-1} \end{aligned}$$

を得て、式 (A. 51) を考慮すると、命題 a. 5 の(4)より、不等式

$$g(y_i) \cdot [1 - g(y_i)] \leq g(y_i^-) \cdot [1 - g(y_i^-)] \quad (\text{A. 58})$$

が成り立つ。ここで、式 (A. 20) を考慮すると、

$$\begin{aligned} & \exp[-c \cdot g^{-1}(x_i^*)] \\ &= \exp[-\log_e [p^{-1}(1-p)]] \\ &= p/(1-p) \end{aligned} \quad (\text{A. 59})$$

であるから、式 (A. 48) を考慮すると、

$$\exp[-cy_i^-]$$

$$\begin{aligned}
&= \exp[-c \cdot g^{-1}(x_i^*)] \cdot \exp(c\delta \|W_i\|) \\
&= [p/(1-p)] \cdot \exp[c\delta \|W_i\|] \tag{A. 60}
\end{aligned}$$

を得て、命題 a. 3 から、

$$\begin{aligned}
&g(y_i^-) \cdot [1 - g(y_i^-)] \\
&= \frac{1}{1 + p(1-p)^{-1} \cdot \exp[c\delta \|W_i\|]} \cdot \frac{p(1-p)^{-1} \cdot \exp[c\delta \|W_i\|]}{1 + p(1-p)^{-1} \cdot \exp[c \cdot \delta \|W_i\|]} \\
&= \frac{(1-p) \cdot \exp[-c\delta \|W_i\|]}{(1-p) \exp[-c\delta \|W_i\|] + p} \cdot \frac{p}{(1-p) \exp[-c\delta \|W_i\|] + p} \\
&= \frac{p(1-p)}{(1-p) \exp[-c\delta \|W_i\|] + p} \cdot \frac{\exp[-c\delta \|W_i\|]}{(1-p) \exp[-c\delta \|W_i\|] + p} \\
&= \frac{p(1-p)}{p + (1-p) \cdot \exp[-c\delta \|W_i\|]} \cdot \frac{1}{p \exp[c\delta \|W_i\|] + (1-p)} \\
&= \frac{p(1-p)}{p + (1-p) \cdot \exp[-c\delta \|W_i\|]} \cdot \frac{1}{1 + p[\exp[c \cdot \delta \|W_i\|] - 1]} \\
&\leq \frac{p(1-p)}{p + (1-p) \cdot \exp[-c\delta \|W_i\|]} \\
&\leq p(1-p) \cdot \exp[c\delta \|W_i\|] \quad \because \text{命題 a. 4}
\end{aligned}$$

つまり、

$$g(y_i^-) \cdot [1 - g(y_i^-)] \leq p(1-p) \exp(c\delta \|W_i\|) \tag{A. 61}$$

を得て、2式 (A. 58), (A. 61) から、式 (A. 53) の成立がわかる。

さて、この (A. 53) を使い、不等式 (A. 38) を証明しよう。

式 (A. 52) より

$$\begin{aligned}
&\exp[c\delta \|W_i\|] \\
&\leq \exp[\rho \log_e [p^{-1}(1-p)]] \\
&= [p^{-1}(1-p)]^\rho
\end{aligned}$$

であり、式 (A. 21) を考慮すると、

$$\|W_i\| = [c^{-1} \log_e [p^{-1}(1-p)]] \cdot \|V_i\|$$

であるから、2式 (A. 45), (A. 53) より

$$\begin{aligned}
&|G(x) - G_i(z)| \\
&\leq c \cdot p(1-p) \cdot \exp[c\delta \|W_i\|] \cdot \|W_i\| \cdot \|\underline{x} - \underline{z}\| \\
&\leq cp(1-p) [p^{-1}(1-p)]^\rho \cdot [c^{-1} \log_e [p^{-1}(1-p)]] \cdot \|V_i\| \cdot \|\underline{x} - \underline{z}\| \\
&\leq p(1/p)^\rho \cdot [\log_e p^{-1}] \cdot \|V_i\| \cdot \|\underline{x} - \underline{z}\| \\
&\quad \because \text{式 (A. 35), 式 (A. 19)} \\
&\leq p \cdot p^{-\rho} [\log_e p^{-1}] (\rho/\delta) \cdot \|\underline{x} - \underline{z}\| \quad \because \text{式 (A. 35)}
\end{aligned}$$

$$\begin{aligned} &\leq (\rho/\delta) \cdot p^{1-\rho} [\log_e p^{-1}] \cdot \|\underline{x} - \underline{z}\| \\ &= \kappa \cdot \|\underline{x} - \underline{z}\| \quad \therefore \quad (\text{A. 39}) \end{aligned}$$

の成立がわかり、式 (A. 38) の証明が終了した。

(系の証明)

$$\lceil \lim_{x \rightarrow a+0} f_1(x) = \lim_{x \rightarrow a+0} f_2(x) = +\infty$$

のとき、

$$\lim_{x \rightarrow a+0} [df_1(x)/dx] / [df_2(x)/dx] = \ell$$

ならば

$$\lim_{x \rightarrow a+0} f_1(x)/f_2(x) = \ell \text{ が成り立つ}$$

を適用すれば、

$$\begin{aligned} &\lim_{p \rightarrow a+0} \kappa \\ &= -(\sqrt{n}\rho/\delta) \cdot \lim_{p \rightarrow +0} p^{1-\rho} \log_e p \\ &= -(\sqrt{n}\rho/\delta) \cdot \lim_{p \rightarrow +0} (\log_e p) / p^{\rho-1} \\ &= -(\sqrt{n}\rho/\delta) \cdot \lim_{p \rightarrow +0} p^{-1} / [(\rho-1)p^{\rho-2}] \\ &= -(\sqrt{n}\rho/\delta) \cdot [1/(\rho-1)] \cdot \lim_{p \rightarrow +0} (1/p^{\rho-1}) \\ &= -(\sqrt{n}\rho/\delta) \cdot [1/(\rho-1)] \cdot \lim_{p \rightarrow +0} p^{1-\rho} \\ &= 0 \quad \therefore \text{式 (A. 35)} \end{aligned}$$

を得て、証明された。 □

A. 3 想起過程の一意的収束性

n 次元ユークリッド空間 R^n の部分集合 X は $\underline{x}, \underline{z} \in X$ かつ $1 \geq \alpha \geq 0$ ならば

$$\alpha \cdot \underline{x} + (1-\alpha) \cdot \underline{z} \in X$$

を満足するとき、凸集合 (convex set) であるといわれる。

A. 2 節でのニューラルネットが、ある条件の下で唯一つの解を持つことが次の定理 A. 3 で指摘される。

[定理 A. 3] (想起過程の一意的収束定理)

記憶パターンの集合

$$\underline{x}^a = (x_1^a, x_2^a, \dots, x_n^a), \quad a = 1 \sim m$$

内のいずれか1つを

$$\underline{x}^* = (x_1^*, x_2^*, \dots, x_n^*) \in R^n$$

と表記しよう。条件

$$a \neq b \text{ のとき, } \|\underline{x}^a - \underline{x}^b\| > 0 \quad (\text{A. 62})$$

を設定する。また、式 (A. 19) での正数 p を、定理 A. 2 の系が成立するように、十分小

さく選んでいるものとし、式 (A. 39) の κ を

$$0 < \kappa < 1 \quad (\text{A. 63})$$

とする。

(i) 定理 A. 2 での半径 δ の球 $B_\delta(\underline{x}^*)$ は凸集合である。

(ii) $\underline{x} \in B_\delta(\underline{x}^*)$ のとき、式 (A. 29) の写像 G について、 $G(\underline{x}) \in B_\delta(\underline{x}^*)$ が成り立つ。

(iii) 不動点方程式

$$\underline{x} = G(\underline{x}) \quad (\text{A. 64})$$

は、球 $B_\delta(\underline{x}^*)$ の中に唯一つの解 \underline{x}^* を持ち、この解 \underline{x}^* は、
入力パターン \underline{x} が

$$\underline{x} \in B_\delta(\underline{x}^*) \quad (\text{A. 65})$$

を満たしていれば、

$$\underline{x}(t) |_{t=0} = \underline{x} \in B_\delta(\underline{x}^*) \quad (\text{A. 66})$$

として、生成されるパターン列

$$\underline{x}(t+1) = G(\underline{x}(t)), t = 0, 1, 2, \dots \quad (\text{A. 67})$$

の極限として

一意的に得られる。つまり、パターン列

$$\underline{x}(t), t = 0, 1, 2, \dots \quad (\text{A. 68})$$

は一意的に \underline{x}^* に収束し、 \underline{x}^* 以外には以束しない。

備考： $\underline{x}(t) |_{t=0} = \underline{x}$ と設定された処理対象パターン $\underline{x} = (x_1 x_2 \dots x_n)$ の各成分 x_i は $p, 1-p$ 以外の実数から成っていてもかまわない。この定理 A. 3 によれば、

入力パターン \underline{x} が球 $B_\delta(\underline{x}^a)$ の和集合

$\cup_{a=1}^m B_\delta(\underline{x}^a)$ に属していれば、記憶パターンの1つ \underline{x}^a として想起認識される

ということになる。しかしながら、入力パターン \underline{x} がこの和集合 $\cup_{a=1}^m B_\delta(\underline{x}^a)$ の外にある場合は、記憶パターンの1つ \underline{x}^a に収束することが保証されず、この入力パターン \underline{x} を出発点として、反復式 (A. 67) が実行されても、 $\cup_{a=1}^m B_\delta(\underline{x}^a)$ の外にあるパターン \underline{z} に収束するかも知れない。□

(iv) 球 $B_\delta(\underline{x}^a)$, $a = 1 \sim m$ は互いに素である。

(v) 定数 τ を

$$\tau \equiv \min_{a,b \in \{1,2,\dots,m\}, a \neq b} \|\underline{x}^a - \underline{x}^b\| \quad (\text{A. 69})$$

と定義すると、定数 δ が不等式

$$\delta < 2^{-1} \cdot \tau \quad (\text{A. 70})$$

を満たしていれば、球 $B_\delta(\underline{x}^a)$, $a = 1 \sim m$ は互いに素である。つまり、

$$B_\delta(\underline{x}^a) \cap B_\delta(\underline{x}^b) = \phi(a \neq b). \quad (\text{A. 71})$$

(証明)

i の証明 : $\underline{x}, \underline{z} \in B_\delta(\underline{x}^*), 0 \leq \alpha \leq 1$

について,

$$\begin{aligned} & \| \alpha \underline{x} + (1 - \alpha) \underline{z} - \underline{x}^* \| \\ &= \| \alpha(\underline{x} - \underline{x}^*) + (1 - \alpha)(\underline{z} - \underline{x}^*) \| \\ &\leq \alpha \| \underline{x} - \underline{x}^* \| + (1 - \alpha) \| \underline{z} - \underline{x}^* \| \\ &\leq \alpha \cdot \delta + (1 - \alpha) \cdot \delta = \delta \end{aligned}$$

を得,

$$\underline{dx} + (1 - \alpha) \underline{z} \in B_\delta(\underline{x}^*).$$

ii の証明 : $\underline{x} \in B_\delta(\underline{x}^*)$ のとき, $\| \underline{x} - \underline{x}^* \| \leq \delta$ であるから, 不動点方程式

$$G(\underline{x}^*) = \underline{x}^*$$

に, 定理 A. 2 の式 (A. 38) を適用して, 定理 A. 2 の系あるいは式 (A. 63) を考慮すれば,

$$\begin{aligned} & \| G(\underline{x}) - \underline{x}^* \| = \| G(\underline{x}) - G(\underline{x}^*) \| \\ &\leq \kappa \cdot \| \underline{x} - \underline{x}^* \| \leq \kappa \cdot \delta < \delta \\ &\therefore G(\underline{x}) \in B_\delta(\underline{x}^*) \end{aligned} \quad (\text{A. 72})$$

iii の証明 : 式 (A. 66) の $\underline{x}(0)$ に ii を適用して,

$$\underline{x}(t) \in B_\delta(\underline{x}^*), t = 0, 1, 2, \dots \quad (\text{A. 73})$$

がいえ, 定理 A. 2 の式 (A. 38) を適用して,

$$\begin{aligned} & \| \underline{x}(t+1) - \underline{x}(t) \| \\ &= \| G(\underline{x}(t)) - G(\underline{x}(t+1)) \| \\ &\leq \kappa \cdot \| \underline{x}(t) - \underline{x}(t-1) \| \\ &\leq \dots \\ &\leq \kappa^{t-1} \cdot \| \underline{x}(2) - \underline{x}(1) \|, t = 1, 2, \dots \end{aligned}$$

が得られる。これから,

$s > t$ のとき,

$$\begin{aligned} & \| \underline{x}(s) - \underline{x}(t) \| \\ &\leq \sum_{u=t}^{s-1} \| \underline{x}(u+1) - \underline{x}(u) \| \\ &\leq \kappa^{t-1} \cdot [\kappa^{s-t-1} + \dots + \kappa + 1] \cdot \| \underline{x}(2) - \underline{x}(1) \| \\ &\leq \kappa^{t-1} \cdot [1 - \kappa]^{-1} \cdot \| \underline{x}(2) - \underline{x}(1) \| \\ &\rightarrow 0 \quad (t \rightarrow \infty) \end{aligned} \quad (\text{A. 74})$$

の成立がわかる。よって, R^n がノルム $\| \cdot \|$ に関し完備であるから, 4 式 (A. 65) ~ (A.

68) でのパターン列 $\{\underline{x}(t)\}_{t=0,1,2,\dots}$ は極限 $\underline{z} \in R^n$ を持ち、

$$\lim_{t \rightarrow \infty} \|\underline{x}(t) - \underline{z}\| = 0 \quad (\text{A. 75})$$

が成り立つ。ところで、4式 (A. 27) ~ (A. 30) での写像 G の定義域を半径 δ の球 $B_\delta(\underline{x}^*)$ に制限したとき、 $B_\delta(\underline{x}^*)$ は閉集合であり、iiより

G の値域 $\subset B_\delta(\underline{x}^*)$ であるから $\underline{z} \in B_\delta(\underline{x}^*)$ である。

式 (A. 73) に定理 A. 2 を適用して、式 (A. 75) から、

$$\begin{aligned} & \|G(\underline{x}(t)) - G(\underline{z})\| \\ & \leq \kappa \cdot \|\underline{x}(t) - \underline{z}\| \rightarrow 0 \quad (t \rightarrow \infty) \end{aligned}$$

の成立を得、よって、等式 (A. 67) において極限に移れば、左辺は \underline{z} に、右辺は $G(\underline{z})$ に移行するから、不動点方程式

$$\underline{z} = G(\underline{z}) \quad (\text{A. 76})$$

が成り立つ。その実、

$$\underline{z} = \underline{x}^* \quad (\text{A. 77})$$

である。以下、その証明。

式 (A. 73) が成立しているから、式 (A. 76) に注意し、定理 A. 2 を適用して、

$$\begin{aligned} \|\underline{x}(t+1) - \underline{x}^*\| &= \|G(\underline{x}(t)) - G(\underline{x}^*)\| \\ &\leq \kappa \cdot \|\underline{x}(t) - \underline{x}^*\| \end{aligned}$$

を得るから、 $t \rightarrow \infty$ に移行すれば、

$$\begin{aligned} \|\underline{z} - \underline{x}^*\| &\leq \kappa \cdot \|\underline{z} - \underline{x}^*\| \\ \therefore (1 - \kappa) \cdot \|\underline{z} - \underline{x}^*\| &\leq 0 \end{aligned}$$

が成立することがわかる。 $\kappa < 1$ であるから、 $\|\underline{z} - \underline{x}^*\| = 0$ を得、証明された。

不動点方程式 $\underline{x} = G(\underline{x})$ の解が一意的である

ことを証明すれば、iiiの証明は完了する。

$\underline{y} \in B_\delta(\underline{x}^*)$ を別の解として、 $\underline{y} = G(\underline{y})$ をする。式 (A. 76) の \underline{z} について、 $\underline{y} = \underline{z}$ を示さなければならない。定理 A. 2 を適用して、

$$\begin{aligned} \|\underline{z} - \underline{y}\| &= \|G(\underline{z}) - G(\underline{y})\| \\ &\leq \kappa \cdot \|\underline{z} - \underline{y}\| \quad \therefore (1 - \kappa) \cdot \|\underline{z} - \underline{y}\| \leq 0 \end{aligned}$$

を得、 $\kappa < 1$ であるから、 $\|\underline{z} - \underline{y}\| = 0$ を得て、証明が終った。

ivの証明： $a \neq b$ とする。 $\|\underline{x}^a - \underline{x}^b\| > 0$ 、つまり $\underline{x}^a \neq \underline{x}^b$ に注意しておく。

球の系 $B_\delta(\underline{x}^a)$ 、 $a = 1 \sim m$ は互いに素でない

と仮定しよう。そうすれば、

$$\underline{x} \in B_\delta(\underline{x}^a) \cap B_\delta(\underline{x}^b)$$

を満たす \underline{x} が存在する。ところが, $\underline{x} \in B_\delta(\underline{x}^a)$ であるから, 2式 (A. 66), (A. 67) に iii を適用して, \underline{x}^a に一意的に収束する。また, $\underline{x} \in B_\delta(\underline{x}^b)$ であるから, 2式 (A. 66), (A. 67) に iii を適用して, \underline{x}^b に一意的に収束する。

収束の一意性から, $\underline{x}^a = \underline{x}^b$ を得, これは矛盾である。

v の証明: 結論を否定すれば,

$$\underline{x} \in B_\delta(\underline{x}^a) \cap B_\delta(\underline{x}^b) \quad (a \neq b)$$

を満たす $\underline{x} \in R^n$ が存在する。

$$\underline{x} \in B_\delta(\underline{x}^a) \quad \text{ということから} \quad \|\underline{x} - \underline{x}^a\| \leq \delta < 2^{-1}\tau$$

$$\underline{x} \in B_\delta(\underline{x}^b) \quad \text{ということから} \quad \|\underline{x} - \underline{x}^b\| \leq \delta < 2^{-1}\tau$$

であるが, ノルム $\|\cdot\|$ の三角不等式, 対称性より,

$$\begin{aligned} \|\underline{x}^a - \underline{x}^b\| &\leq \|\underline{x}^a - \underline{x}\| + \|\underline{x} - \underline{x}^b\| \\ &= \|\underline{x}^a - \underline{x}\| + \|\underline{x}^b - \underline{x}\| \\ &\leq 2\delta < \tau = \min_{a,b \in \{1,2,\dots,m\}, a \neq b} \|\underline{x}^a - \underline{x}^b\| \end{aligned}$$

という矛盾が得られた。 □

付録B (不動点方程式の一解法⁽²⁾)

不動点方程式 (fixed-point equation)

$$\underline{x}^* = F(\underline{x}^*, 0) \tag{B. 1}$$

を解くことを考えよう。ここに, F は

$$F: R^n \times [0, \infty) \rightarrow R^n \tag{B. 2}$$

ここに, R^n は n 次元ユークリッド空間であり,

$$[0, \infty) \equiv \{\varepsilon \mid 0 \leq \varepsilon < \infty\}$$

という写像である。この不動点方程式 (B. 1) を an iterative scheme

初期条件

$$\underline{x} = \underline{x}_t \mid_{t=0} \in R^n \tag{B. 3}$$

の下で,

$$\begin{aligned} \underline{x}_{t+1} &= F(\underline{x}_t, \varepsilon_t) \in R^n, \\ \varepsilon_t &\in [0, \infty), t = 0, 1, 2, \dots, N-1 \end{aligned} \tag{B. 4}$$

で解くことを考える。ただし, 2条件

$$(i) \quad \forall \underline{x} \in R^n, \|F(\underline{x}, \varepsilon) - F(\underline{x}, 0)\| \leq \varepsilon$$

ここに, $\|\cdot\|$ は固定したノルム

(ii) $\exists L$ (Lipschitz constant) ($0 < L < 1$),

$$\forall \underline{x}, \forall \underline{y} \in R^n,$$

$$\|F(\underline{x}, 0) - F(\underline{y}, 0)\| \leq L \cdot \|\underline{x} - \underline{y}\|$$

を課しておく。なお、以下の the contraction-mapping theorem (縮小写像定理) によれば⁽²²⁾, R^n は距離

$$\text{dis}(\underline{x}, \underline{y}) \equiv \|\underline{x} - \underline{y}\| = [\sum_{j=1}^n (x_j - y_j)^2]^{1/2}$$

$$\text{ここに, } \underline{x} = {}^t(x_1 \ x_2 \ \cdots \ x_n)$$

$$\underline{y} = {}^t(y_1 \ y_2 \ \cdots \ y_n)$$

に関して完備であるから、条件 ii の下で、the traditional iteration

$$\underline{x}_{t+1} = F(\underline{x}_t, 0) \tag{B. 5}$$

は唯一つの \underline{x}^* に収束して、不動点方程式 (B. 1) の成り立つ \underline{x}^* が求められることが判明している。

ここに、 t は転置の意である。

〔縮小写像の原理〕

作用素 A が完備な距離空間 X を自分自身の中に写し、その定義域 Domain (A) が Domain (A) = X であり、かつ α を条件

$$0 < \alpha < 1$$

を満たす数として、 X の中の任意の φ, η に対して、

$$\text{dis}(A\varphi, A\eta) \leq \alpha \cdot \text{dis}(\varphi, \eta)$$

あるならば、不動点方程式

$$\varphi = A\varphi$$

は唯一つの解 φ^* を持つ。この解 φ^* は、出発の元 φ_1 を任意に選んでできる列

$$\varphi_2 = A\varphi_1, \varphi_3 = A\varphi_2, \varphi_4 = A\varphi_3, \dots$$

の極限として得られる。ここに、 $\text{dis}(\varphi, \eta)$ は X の上での距離関数である。 □

さて、 $\|\underline{x}_N - \underline{x}^*\|$ が

$$\begin{aligned} & \|\underline{x}_N - \underline{x}^*\| \\ & \leq \sum_{i=0}^{N-1} L^{N-1-i} \cdot \varepsilon_i + L^N \cdot \|\underline{x}_0 - \underline{x}^*\| \end{aligned} \tag{B. 6}$$

と評価されることに留意して、

$$\sum_{i=0}^{N-1} L^{N-1-i} \cdot \varepsilon_i + L^N \cdot \|\underline{x}_0 - \underline{x}^*\| = a \text{ (与えられた正数)} \tag{B. 7}$$

を条件として、

$$\sum_{t=0}^{N-1} \text{cost}(\varepsilon_t) \rightarrow \min \quad (\text{B. 8})$$

ならしめる正整数 N と制御ベクトル

$$e = {}^t(\varepsilon_0, \varepsilon_1, \dots, \varepsilon_{N-1})$$

とを見つけよ (COP; the constrained minimization problem)

という問題に転換して考えれば, その解答は次のように与えられる。

〔定理 B. 1〕 (指数関数形コストの COP 解定理)

ある $p > 0$ に対し,

$$\text{cost}(\varepsilon) = \varepsilon^{-p} \quad (\text{B. 9})$$

と設定すれば, COP の解を与える N と ε_t は

$$N = -(p+1) \cdot \frac{\log_e [\|\underline{x}_0 - \underline{x}^*\|/a]}{\log_e L} \quad (\text{B. 10})$$

$$\varepsilon_t = \kappa \cdot L^{(1+t-N)/(p+1)}, t = 0, 1, 2, \dots, N-1 \quad (\text{B. 11})$$

で与えられる。ここに,

$$\kappa = \frac{(1-\hat{L})}{(1-\hat{L}^N)} \cdot [a - L^N \cdot \|\underline{x}^* - \underline{x}_0\|] \quad (\text{B. 12})$$

$$\hat{L} = L^{p/(p+1)} \quad (\text{B. 13})$$

特に, ノルム $\|\cdot\|$ を

$$\|\underline{x} - \underline{y}\|^2 = \sum_{j=1}^n |x_j - y_j|^2 \quad (\text{B. 14})$$

と設定しているならば, 条件 i は,

$$\sum_{j=1}^n v_j^2 \leq 1 \quad (\text{B. 15})$$

を満たす $v = {}^t(v_1, v_2, \dots, v_n)$ を用意し,

$$F_j(\underline{x}, \varepsilon) = F_j(\underline{x}, 0) + \varepsilon \cdot v_j, j = 1, 2, \dots, n \quad (\text{B. 16})$$

と設定すれば, 満たされる。ここに, 式 (B. 2) の $F(\underline{x}, \varepsilon)$ は, t を転置の意として,

$$F(\underline{x}, \varepsilon) = {}^t(F_1(\underline{x}, \varepsilon), F_2(\underline{x}, \varepsilon), \dots, F_n(\underline{x}, \varepsilon)) \quad (\text{B. 17})$$

である。

付録 C (モデル構成作用素 $T: \Phi \rightarrow \Phi$ の構成例)

$T\phi \in \Phi$ がパターン $\phi \in \Phi$ の代りとなるモデルであるとすれば, S. Suzuki の提案している “パターン認識の数学的理論⁽²⁰⁾” では, 写像

$$T: \Phi \rightarrow \Phi \tag{C. 1}$$

は次の4性質 イ～ニ を満たさなければならない⁽²¹⁾。

[モデル構成作用素 T の満たすべき4性質]

(性質イ) $\varphi = 0 \in \Phi$ について $T\varphi = \varphi$.

(性質ロ) $\forall \varphi \in \Phi, T(a\varphi) = T\varphi$ for any positive real number a .

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

(性質ニ) $\exists \varphi \in \Phi, T\varphi \neq 0$.

□

$0 \in \Phi \subset$ Hilbert space \mathfrak{H} であつ

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

を満たすパターン集合 Φ

を選定する。つまり、 Φ は

$$\varphi \in \Phi \text{ ならば必ず } T\varphi \in \Phi$$

を満たし、 Φ は写像 T に関する不変集合であらねばならない。

Invariant pattern recognition with high discrimination and robustness to noise

の確保のために、パターンの変形に対し不変な特徴を抽出する技術 (a technique to extract invariants for deformations) が必要であり、これらについては文献(24), (25)にあるが、この特徴抽出技術を用いれば、少くともある1つの1パラメータ Lie 座標変換群の下で不変な“パターン φ のモデル $T\varphi$ ” を構成できる。

本付録Cでは、ユニタリ座標変換不変性を備えていないが、上記の4性質 イ～ニ を満たす有用なパターンモデル $T\varphi$ を1つ指摘しておこう。ユニタリ座標変換不変性を備え、4性質 イ～ニ を満たすパターンモデル $T\varphi$ の諸例については、文献(3), (7)~(10), (11), (13), (16)~(18), (20), (21)にある。

パターン $\varphi_k \in \mathfrak{H}$ の1次独立な系

$$\varphi_1, \varphi_2, \dots, \varphi_k, \dots \tag{C. 2}$$

を選定し、その後、

$$\begin{aligned} \tilde{\varphi}_1 &= \varphi_1 \\ \tilde{\varphi}_2 &= \varphi_2 - (\varphi_2, \tilde{\varphi}_1 \| \tilde{\varphi}_1 \|^{-1}) \cdot \tilde{\varphi}_1 \| \tilde{\varphi}_1 \|^{-1} \\ &\vdots \\ \tilde{\varphi}_k &= \varphi_k - \sum_{j=1}^{k-1} (\varphi_k, \tilde{\varphi}_j \| \tilde{\varphi}_j \|^{-1}) \cdot \tilde{\varphi}_j \| \tilde{\varphi}_j \|^{-1} \\ &\vdots \end{aligned} \tag{C. 3}$$

(Gram-Schmidt Orthogonalization process)

を求め、

$$\eta_k \equiv \tilde{\varphi}_k, k = 1, 2, \dots \tag{C. 4}$$

とおいて得られる \mathfrak{H} の元から成る系 $\{\eta_k\}_{k \in L}$ は

$$\forall k, \forall \ell \in L = \{1, 2, \dots\},$$

$$(\eta_k, \eta_\ell) = \|\eta_k\| \cdot \|\eta_\ell\| \cdot \delta_{k\ell} \quad (\text{orthogonality})$$

ここに, $\delta_{k\ell} = 1$ if $k = \ell$, $= 0$ if $k \neq \ell$ (C. 5)

を満たす。ここで, 正定数 $C > 0$ を用意し,

$$|c_k| = \sqrt{C} \cdot \|\eta_k\|^{-1} \quad (\text{C. 6})$$

と定義される複素定数 c_k を考え,

$$\phi_k \equiv c_k \cdot \eta_k \quad (\text{C. 7})$$

とおけば, $\{\phi_k\}_{k \in L}$ は

$$\forall k \in L, \|\phi_k\|^2 = C(k \in L \text{ に無関係な定数}) > 0$$

(a flat-power property) (C. 8)

が満たされている直交系である。

さて,

$$\text{各 } (\varphi, \phi_k) \text{ は実数値である} \quad (\text{C. 9})$$

とし, 各閾値 e_k, e'_k が不等式

$$0 \leq e_k, e'_k < \|\phi_k\|^2 / \sup_{\ell \in L} \|\phi_\ell\|^2 \quad (\text{C. 10})$$

を満たすように選ばれているとしよう。このとき,

$$\hat{a}_k(\varphi) \equiv \begin{cases} 0 \cdots \forall k \in L, (\varphi, \phi_k) = 0 & \text{の場合} \\ (\varphi, \phi_k) / \sup_{k \in L} |(\varphi, \phi_k)| \cdots \exists k \in L, (\varphi, \phi_k) \neq 0 & \text{の場合} \end{cases} \quad (\text{C. 11})$$

として,

$$u(\varphi, k) \equiv \begin{cases} +1 \cdots + e_k < \hat{a}_k(\varphi) \leq +1 & \text{のとき} \\ 0 \cdots - e'_k \leq \hat{a}_k(\varphi) \leq + e_k & \text{のとき} \\ -1 \cdots -1 \leq \hat{a}_k(\varphi) < - e'_k & \text{のとき} \end{cases} \quad (\text{C. 12})$$

を用意し,

$$T\varphi \equiv \sum_{k \in L} u(\varphi, k) \cdot \phi_k \quad (\text{C. 13})$$

と定義される写像 $T: \Phi \rightarrow \Phi$ は上記の4性質 $I \sim II$ を満たすことが示され, モデル構成作用素である。

4性質 $I \sim II$ を満たし, しかも 可分な Hilbert 空間 \mathfrak{H} でのあるユニタリ作用素のつくる群の下で不変なパターンモデル $T\varphi$ の構成については, パターン $\varphi \in \Phi \subset \mathfrak{H}$ から抽出される第 $k \in L$ 番目の特徴量 $u(\varphi, k)$ として, ある自己共役作用素 H の関数^{(4), (5), (14), (15)}を用いたパターン φ の2次非負汎関数値 (S. Suzuki の提案した測度的ユニタリ不変量^{(3), (6), (9), (20), (32)} (metrically unitary invariants)) を用いれば可能であり, この構成などについては,

文献(3), (7)~(11), (13), (16), (17), (20), (21)にある。

(鈴木昇一・佐久間拓也, 文教大学情報学部“情報研究 No. 15”投稿論文, パターンモデルを用いた不動点探索形連想記憶システム方程式, 投稿年月日 1994年10月13日)