

修士論文

神経回路網における疎符号化の研究

指導教官 甘利 俊一 教授

平成2年2月

東京大学大学院
工学系研究科 計数工学専攻

赤穂 昭太郎

目次

第1章	序論.....	1
第2章	連想記憶における疎符号化.....	4
§1	パターンの疎符号化.....	4
§2	連想記憶モデル.....	6
§3	疎符号化された連想記憶モデルの容量.....	8
(1)	記憶容量.....	8
(2)	情報容量.....	15
§4	引き込み領域の広さと容量.....	17
第3章	3層回路の記憶容量について.....	22
§1	神経回路網の記憶容量と学習可能性.....	22
§2	2層回路の容量.....	23
§3	3層回路の容量の上限.....	28
§4	3層回路の容量の下限.....	31
(1)	中間層を疎符号化したモデル.....	31
(2)	モデルの容量.....	33
(3)	モデルの意味.....	40

第4章	競合機構をもつ連想記憶モデル	42
§1	基本競合系	42
§2	競合機構と連想記憶モデルの疎符号化	47
第5章	結論	54
謝辞	56
参考文献	57

第1章 序論

人間の脳は高度で複雑な情報処理システムであり、そのメカニズムはまだよくわかっていないというのが実情である。脳が行っている情報処理の特徴を、特に記憶という観点からいくつか列挙してみると、まず脳では個々のニューロンが互いに相互作用をしながら並列的に動作をすることが挙げられる。従って記憶のアクセスなどを高速に行うことができる。またこの記憶の方式はある一つの記憶事項が一つの領域に格納されているのではなく、数多くの記憶が広い領域に互いに重なりあいながら分散的に記憶されていて、多少の損傷に対しても強いというメリットがある。ここで記憶に伴う学習について考えてみると、脳は外界から様々な知識を得て、シナプス荷重を適応的に変化させてその知識を蓄えるとともに、さらにその知識を含むより一般的な法則を学びとっていく。脳は未知の与えられた事物(言葉や画像など)に対して反応し、何らかの処理結果を出力するが、得た知識が少ないうちは一般的な法則を十分に学びとっていないために正しい答を出さ

ことができない。従ってどれだけの知識を例題として与えれば十分な一般的法則を学びとることができるかを明らかにしておく必要がある。これは「帰納的学習における一般化能力」という問題で、これについては神経回路網を含むいろいろな枠組から最近活発に議論され始めた。与えるべき例題の数はモデルの記憶容量、つまりモデルがどれだけのパターンを分類する能力があるかに関係していて、記憶容量が増えれば「増えるほど」与えるべき例題の数を増大させなければならない。

本論文では主として脳の記憶モデルの記憶容量について数理的な解析を行う。その議論の中で中心的な役割を果たすのはパターンベクトルのスパース化(疎符号化)である。スパースなベクトルとはその成分のほとんどが0であって、1である成分の数が非常に少ないものをいう。スパースなパターンベクトルの存在は海馬(記憶の制御機能を果たしていると考えられている脳の部位)における生理学的知見からも得られており、また数理的にもスパースなパターンベクトルを用いることにより記憶容量が飛躍的に増大することが示されている。

まず第2章ではスパース化した連想記憶モデルの容量を求め、容量の増大に寄与するのは出力パターンへのスパース化であることを示す。

第3章では3層の神経回路網の容量について考える。これまでに2層回路の容量は求められており、3層回路についてもその容量の上限だけは求められていた。ここでは3層回路の容量の下限を与えるモデルを示す。このモデルの基本的なアイデアは中間層のパターンをスパース化し、入力パターンをグループ化することによって中間層と出力層との間で記憶するパターンの数を減らすことである。

第4章では競合機構をもったスパースなパターンの連想記憶モデルについて考察する。まず最初に基礎となる競合機構の性質について述べ、その機構にどうやって連想記憶モデルの機構を付加できるかを考察する。

第2章 連想記憶における疎符号化

§1 パターンの疎符号化

神経細胞はその細胞への入力の重み付きの和が、あるしきい値を
超えるかどうかで発火するしないが決まる 2値出力の「しきい値素子」と
してモデル化することができる(図2.1)。つまり x を入力, y を出力
として,

$$y = f\left(\sum_i w_i x_i - \theta\right) \quad (2.1)$$

と書ける。ここで w_i はシナプス荷重, θ はしきい値を表し,
 $f(u)$ は u の正負によって 2値を出力する階段関数である。

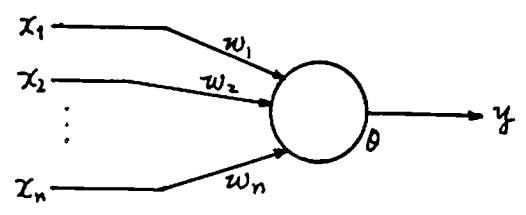


図2.1 神経細胞の工学モデル

この素子を用いて, 入力層 n 次元, 出力層 l 次元の 2層の回路
網を作る(図2.2)。するとその入出力関係は,

$$y_i = f\left(\sum_j w_{ij} x_j - \theta_i\right) \quad (2.2)$$

となり, この変換を $y = T x$ と書くことにする。

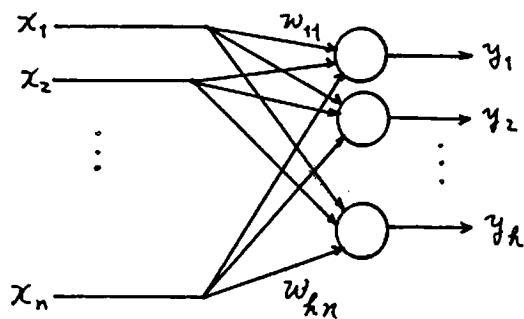


図2.2 2層の神経回路網

さてここで「スパースなパターンベクトル(疎符号)」というものを考える。従来は通常、各成分が 0 か 1 (または -1 か 1) を確率 $\frac{1}{2}$ ずつでとるランダムベクトルを考えたが、ここでは 1 をとる成分の数が非常に少なく限定されているモデルを考える。つまり各成分はある非常に小さな確率 a で 1 をとり、確率 $1-a$ で 0 をとる。従って漸近的に na 個の成分は 1 をとり $n(1-a)$ 個の成分は 0 をとることになる。こうしてできるベクトルをスパースであるといい、スパースなベクトルは 4 章で述べる適当な競合機構によって実現できる。本論文ではスパースでない場合も含めて次のように定式化する。つまりベクトル x は

$$x_i = \begin{cases} 1-a & \dots \text{確率 } a \\ -a & \dots \text{確率 } 1-a \end{cases} \quad (2.3)$$

で決められるとし、このときベクトル x を a -スパースであるということ

にする。スパースであるものとしては例えば $\alpha = \frac{C \log n}{n}$

(n はベクトルの次元, C は定数) や $\alpha = \frac{C}{n}$ などを考える。

このモデルは $\alpha = \frac{1}{2}$ のときは (正規化すれば) ± 1 モデルとなり,

α が非常に小さい場合には 0-1 モデルに近づくという意味で ± 1

モデルと 0-1 モデルを統合した形をしており, (2.3) 式を一見すると

複雑な仮定に見えるが実は $E[x_i] = 0$ となって解析は逆に

易くなる。また Amari^[4] は 0-1 モデルで解析を行い, $\alpha > O(n^{-\frac{1}{2}})$

の場合に非常に性質が悪くなることを示している。

また (2.2) 式で与えた $f(u)$ の形は出力されるパターンのスパース化の度合いによって次のように書ける。つまり y_i が β -スパースのとき,

$$f(u) = \begin{cases} 1 - \beta & \dots u \geq 0 \\ -\beta & \dots u < 0 \end{cases} \quad (2.4)$$

となる。

§2 連想記憶モデル

連想記憶モデルとは, 入力パターン x^1, \dots, x^m と対応する出力パターン y^1, \dots, y^m が与えられたときに, x^u ($u=1, \dots, m$) また

はそれに近いベクトルを図2.2の回路に入力して y^{μ} が出力されるように(2.2)式の w_{ij} や θ_i をうまく選ぶモデルであり、いろいろなモデルが考えられているがここでは相関学習による連想記憶モデルを考える。^{[6][12][16]}

相関学習とはシナプス荷重を、入力パターンと出力パターンの間の相関に比例して増やしていく、つまり $\Delta w_{ij} \propto y_i^{\mu} x_j^{\mu}$ に従って w_{ij} を変化させていく学習方式である。すべてのパターンが等確率で与えられるとすれば、適当な正規化によって

$$w_{ij} = \frac{1}{n} \sum_{\mu=1}^m y_i^{\mu} x_j^{\mu} \quad (2.5)$$

に収束する。

このモデルは非常に自然で単純なモデルであるので数多くの研究がなされた。特に最近になって物理系におけるスピングラスとの類推から、レプリカ対称性やエネルギー関数などの手法・概念を用いた解析が活発に行われるようになった。^{[10][11]}

またGardner^[9]は自己想起モデルの平衡状態の分布を調べ、記憶されたパターンの極近くに平衡状態の固まりがあることを示した。

§3 疎符号化された連想記憶モデルの容量

(1) 記憶容量

入力パターン x^1, \dots, x^m と、対応する出力パターン y^1, \dots, y^m をランダムに決めたとし、すべての $\mu=1, \dots, m$ について

$$y^\mu = T x^\mu$$

が成立する最大の m を記憶容量という。

相関学習を用いた連想記憶モデルの場合、 x_i^μ および y_i^μ がともに ± 1 を確率 $\frac{1}{2}$ でランダムにとるとき、入出力ベクトルの次元を n として記憶容量は $\frac{n}{2 \log 2}$ となることが知られている。

[13][20] さらに $x^\mu = y^\mu$ となる自己想起モデルの場合には、出力を再び入力にフィードバックしてその平衡状態を出力とするモデルを考えるとその容量は約 $0.15n$ になることが知られている。[3][5][11]

最近になって、記憶するパターンをスパースにすると記憶容量が飛躍的に増大することがわかってきた。[4][21][17]

実際にその容量を計算する前に、スパース化によって容量が増大する理由について定性的に考えてみる。

図 2.1 で示した 1 つの素子を考え、入力として x^m を入れたとき、出力 y^m を出したいとする。ここで y^m は 1 か 0 をとるものとする。この素子は信号空間 $X \subset \mathbb{R}^n$ を $n-1$ 次元超平面で切って、その一方の半平面にある信号には 1 を、他方の半平面にある信号には 0 を対応させるものと見ることができ。従って $\{x^m\}$ と $\{y^m\}$ との対応を記憶するには $y^m=1$ となる x^m の集合と $y^m=0$ となる x^m の集合とが線形分離可能でなければならない。 y^m が 1 と 0 とを確率 $\frac{1}{2}$ でとるとすると線形分離可能である確率は次章で述べるように n が十分大きいとき、 $\frac{m}{n} < 2$ ではほぼ 1 となり、 $\frac{m}{n} > 2$ ではほぼ 0 となる。パターンをスパース化するということは $\{y^m\}$ のうちのほとんどが 0 であることを意味する。従ってほとんどの x^m に対しては $y^m=0$ を出力し、非常に少数の x^m に対してだけ $y^m=1$ を出力してやればよいことになり、直観的には容量の増加が見込まれる。

それでは実際に容量を求める。まず定理 2.1 で非常に一般的な形を与え、系 2.2.1 で具体例を示すことにする。

定理 2.1 入力が n 次元で a -スパースであり, 出力が h 次元で b -スパースのとき, 記憶容量は n , h が十分大きいとき,

$$C = \frac{n}{8 b (1-b) \log(h C)}$$

の解として与えられる (右辺に C 自身を含んでいることに注意).

ただし $C n a^2 b \rightarrow \infty$ ($n \rightarrow \infty$) とする.

証明 m 個のパターンを記憶しているとするとシナプス荷重は $w_{ij} = \frac{1}{n} \sum_{\mu=1}^m y_i^{\mu} x_j^{\mu}$ となる. 入力 x^1 から出力 y^1 が正しく想起されるのは,

$$\begin{aligned} u_i &= \sum_j w_{ij} x_j^1 \\ &= \frac{1}{n} y_i^1 x^1 \cdot x^1 + \frac{1}{n} \sum_{\mu \neq 1} y_i^{\mu} x^{\mu} \cdot x^1 \end{aligned} \quad (2.6)$$

として,

$$y_i^1 = f(u_i - \theta)$$

となる場合である.

$$\frac{1}{n} x^1 \cdot x^1 = a(1-a)$$

より (2.6) 式は

$$u_i = a(1-a) y_i^1 + N_i$$

と書ける。ただし $N_i = \frac{1}{n} \sum_{\mu=1}^m y_i^\mu x_i^\mu x_i^1$ である。

N_i は独立な確率変数の和だから $\sqrt{[nN_i]} \rightarrow \infty$ ($n \rightarrow \infty$)

ならば中心極限定理が成立し、 N_i は漸近的に正規分布

となる。 N_i は平均 0 であり分散は、

$$V[y_i^\mu] = b(1-b),$$

$$V[x_i^\mu x_i^1] = a^2(1-a)^2 \quad (\mu \neq 1, i \neq j)$$

によって、

$$\begin{aligned} V[N_i] &= \frac{m}{n^2} \cdot n \cdot b(1-b) \cdot a^2(1-a)^2 \\ &= \frac{m}{n} b(1-b) a^2(1-a)^2 \end{aligned} \quad (2.7)$$

となり、 $\sigma^2 = V[N_i]$ とおくと $N_i \sim N(0, \sigma^2)$ となる。

従って N_i の条件付分布は図 2.3 のように書ける。

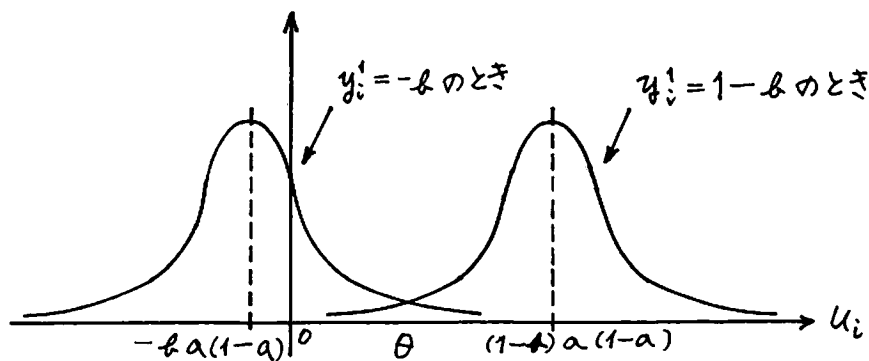


図 2.3 N_i の条件付分布

ここでしきい値 θ を

$$\theta = -b a(1-a) + \frac{1}{2} a(1-a)$$

ととれば $y_i^1 = 1-b$ の場合も $y_i^1 = -b$ の場合も誤りが起きる確率は次のように書ける。

$$p_0 = \Pr(f(u_i - \theta) \neq y_i^1) = F\left(\frac{a(1-a)}{2\sigma}\right).$$

ただし

$$F(u) = \frac{1}{\sqrt{2\pi}} \int_u^\infty e^{-\frac{u^2}{2}} du.$$

従って全てのパターンの中の全成分のうち誤っているものの個数の期待値は,

$$E = mh p_0$$

となる。ここで u が十分大きいとき

$$F(u) \simeq \frac{1}{\sqrt{2\pi} u} e^{-\frac{u^2}{2}} \quad (2.8)$$

となることを用いて,

$$E \simeq \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{a^2(1-a)^2}{8\sigma^2} + \log(mh) + O\left(\log\left(\frac{a(1-a)}{2\sigma}\right)\right)\right\}$$

を得る。 $n, h \rightarrow \infty$ のとき $E \rightarrow 0$ となるのは,

$$\frac{a^2(1-a)^2}{8\sigma^2} > \log(mh)$$

のときに限る。従って (2.7) 式より

$$m < \frac{n}{8h(1-h)\log(hm)}$$

を得る。□

証明の中で正しい値を $y_1^* = 1-h$ のときと $y_1^* = -h$ のときの誤り確率が等しくなるようにとったが、誤り個数の期待値が最小になるようにとると定数倍だけ容量を増やすことができる。

定理 2.1 は記憶容量に入力パターンのスプース度 α が関係しないことを表しており興味深い。つまり記憶容量は入力パターンがスプースかどうかに関わらず出力パターンがスプースになればなるほど容量が増加することがわかった。

ここで定理 2.1 は容量 C について陽な形で書かれていないのでいくつかの n について具体的に容量の値を計算してみる。

系 2.2 入力・出力がともに n 次元で $\frac{1}{2}$ -スプース (つまりスプースでない) とし、 n が十分大きければ

$$C = \frac{n}{4 \log n}$$

となる。また入力・出力がともに n 次元で $\frac{C \log n}{n}$ -スプース

であるとする n が十分大きければ

$$C = \frac{n^2}{24c(\log n)^2}$$

となる. また 入力 n 次元で $\frac{1}{2}$ -スパース (つまりスパースでない),

出力 h 次元で $\frac{c}{h}$ -スパースのとき, n と h が十分大きくて

$h \gg \log \log n$ ならば,

$$C = \frac{nh}{8c \log(nh^2)}$$

となる.

証明 定理 2.1 にそのまま代入し, 微小項を無視すれば

よい. またすべての場合に中心極限定理が成り立つため

の仮定 $Cna^2h \xrightarrow{(n \rightarrow \infty)} \infty$ が成り立つ. 例えは最後の

場合

$$\begin{aligned} C &= \frac{nh}{8c \log(hC)} \\ &= \frac{nh}{8c(\log h + \log(nh) + O(\log \log n))} \\ &\simeq \frac{nh}{8c \log(nh^2)} \end{aligned}$$

であり,

$$Cna^2h = O\left(\frac{n}{\log(nh^2)}\right) \rightarrow \infty \quad (n \rightarrow \infty)$$

が成り立つ. \square

(2) 情報容量

出力パターンのスパーズ化によって記憶容量が増大することは示されたが、スパーズ化によって一つ一つのパターンがもつ情報量は減少する。それでは全体としての情報容量はどれだけになるだろうか。情報容量の定義はいくつか考えられるが、ここでは入出力パターン一つがもつ情報量を I 、記憶容量を C として、

$$I_{all} = I \times C \quad (2.9)$$

で定義する。この量は C 個のパターンをコンピュータメモリに格納するのに必要なビット数であるが、神経回路網の記憶はコンピュータメモリに格納するように記憶されているわけではないので厳密な意味での情報容量というわけではないことに注意しておく。

さて一つ一つのパターンがもつ情報量は、 $\frac{1}{2}$ -スパーズの場合（つまりスパーズでない場合） n ビットである。スパーズであるとすれば一つのベクトルの一つの成分がもつ情報量は

$$H_2(a) = -a \log_2 a - (1-a) \log_2 (1-a)$$

であり、各成分は独立だから一つのパターンがもつ情報量は $n H_2(a)$ ビットとなる。(一つのパターンのもつ情報量は $\log_2 \binom{n}{na}$ と見ることもしできるがこれを Stirling の公式で展開しても $n H_2(a)$ ビットとなる。)

以上のことから情報容量について次のことが言える。

定理 2.3 入力が n 次元で a -スパース, 出力が h 次元で b -スパースのとき, 情報容量は n, h が十分大きいとき (2.9) 式の意味で次式のようになる。

$$I_{all} = C \{ n H_2(a) + h H_2(b) \}.$$

ただし C は記憶容量である。

特に $n=h, a=b=\frac{1}{2}$ のときは

$$I_{all} = \frac{n^2}{2 \log n}$$

となり, $n=h, a=b=\frac{c \log n}{n}$ のときは

$$I_{all} = \frac{n^2}{12}$$

となり, $a=\frac{1}{2}, b=\frac{c}{h}, h \gg \log n$ のときは

$$I_{all} = \frac{n^2 h}{8c \log(nh^2)}$$

となる。

証明 定理 2.1, 系 2.2 の結果から容易に計算できる.

ただし $a = \frac{c \log n}{n}$ のとき,

$$nH_2(a) \simeq c(\log n)^2$$

となることを用いた. \square

定理 2.3 は スパース化が単に記憶容量の増大だけでなく情報容量の増大をも引き起こすことを示している.

§4 引き込み領域の広さと容量

§3 では入力パターン x^{μ} (n 次元, a -スパース) を入力して正しい出力パターン y^{μ} (k 次元, b -スパース) を出力する場合の容量を求めたが, ここではスパースな連想記憶モデルの引き込み領域の広さについて考察する. つまり入力として x^{μ} そのものではなく x^{μ} に近いパターン \tilde{x}^{μ} を入力して y^{μ} を想起することを考えたとき, \tilde{x}^{μ} としては x^{μ} からどれだけ離れたところまでとれるかということについて解析する.

x^{μ} が a -スパース のとき x^{μ} の成分のうち na 個は $1-a$ であり, $n(1-a)$ 個は $-a$ である. ここで x^{μ} に近いパターン \tilde{x}^{μ} として次のようなものを考える.

まず x^{μ} の成分のうち $1-a$ をとるもの (na 個) のうち $na\rho$ 個の成分を $-a$ にかえ, さらに x^{μ} の成分のうち $-a$ をとるもの ($n(1-a)$ 個) のうち naq 個の成分を $1-a$ にかえる. 解析を容易にするために $E[\tilde{x}_i^{\mu}] = 0$ となるようにシフトしてやると, $\tilde{a} = (1-\rho+q)a$ として \tilde{x}_i^{μ} は $1-\tilde{a}$ と $-\tilde{a}$ の2値をとることになる.

こうしてできる任意の \tilde{x}^{μ} から y^{μ} を想起する時の記憶容量は次のようになる.

定理 2.4 定理 2.1 の容量 C とすると, 任意の \tilde{x}^{μ} から y^{μ} を想起する場合の容量は a が十分小さければ,

$$C' = \frac{(1-\rho)^2}{1-\rho+q} C$$

となる.

証明 定理 2.1 と同様に証明できる. \tilde{x}^1 から y^1 が正しく想起されるのは, $w_j = \frac{1}{n} \sum_{\mu} y_i^{\mu} x_j^{\mu}$ より,

$$u_i = \sum_j w_j \tilde{x}_j^1 = \frac{1}{n} y_i^1 x^1 \cdot \tilde{x}^1 + \frac{1}{n} \sum_{\mu \neq 1} y_i^{\mu} x^{\mu} \cdot \tilde{x}^1$$

として

$$y_i^1 = f(u_i - \delta)$$

となる場合である。

$$\frac{1}{n} x_i^1 \tilde{x}_i^1 = a \{1 - a - \rho + a(\rho - \xi)\}$$

より,

$$u_i = a \{1 - a - \rho + a(\rho - \xi)\} y_i^1 + N_i$$

と書ける。ただし $N_i = \frac{1}{n} \sum_{m=1}^m y_i^m x_i^m \tilde{x}_i^1$ 。

定理 2.1 の条件を満たしていれば N_i は中心極限定理により漸近的に正規分布となる。 N_i の平均は 0 であり、分散は、

$$V[x_i^m \tilde{x}_i^1] = a \tilde{a} (1 - a) (1 - \tilde{a})$$

より,

$$V[N_i] = \frac{m}{n} \xi (1 - \xi) \cdot a \tilde{a} (1 - a) (1 - \tilde{a}) = \sigma^2$$

となり、 $N_i \sim N(0, \sigma^2)$ となる。

しきい値は定理 2.1 の証明と同様に、簡単のため誤り確率が y_i^1 の値によらず等しくなるようにとると、誤り確率は

$$p_0 = \Pr(f(u_i - \theta) \neq y_i^1) = F\left(\frac{a \{1 - a - \rho + a(\rho - \xi)\}}{2\sigma}\right)$$

となる。従って全パターン of 全成分のうち誤っているものの個数の期待値は、

$$E = mh p_0$$

となり, (2.8) 式の近似を用いると, $E \rightarrow 0$ ($m, n \rightarrow \infty$) となる必要十分条件は,

$$\frac{a^2 \{1-a-p+a(p-q)\}^2}{8\sigma^2} > \log(hm)$$

となり, σ^2 の値を代入すると,

$$m < \frac{na^2 \{1-a-p+a(p-q)\}^2}{8h(1-h)a\tilde{a} \log(hm)}$$

となる. これを定理 2.1 の結果と比較し, $\hat{a} = (1-p+q)a$ に注意すると,

$$C' = \frac{\{1-a-p+a(p-q)\}^2}{1-p+q} C$$

を得る. ここで a が非常に小さいとすれば定理の式が得られる. \square

定理 2.4 は 1 を入力すべき成分を 0 と入力してしまうことが 0 を入力すべき成分を 1 と入力してしまうことと比べて損失が大きいことを示している. 図 2.4 に p, q, C' のグラフを示す.

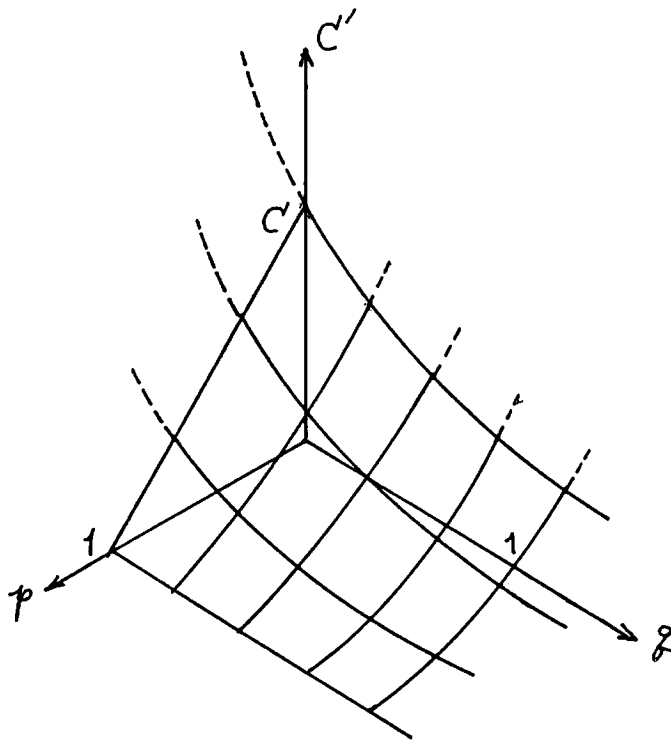


図2.4 引き込み領域の広さと容量の関係

第3章 3層回路の記憶容量について

§1 神経回路網の記憶容量と学習可能性

第2章では連想記憶モデルについて記憶容量を考えたが一般に神経回路網はどれぐらいの記憶容量を持つかというのは興味のある問題である。

また最近になって帰納的学習の学習可能性について活発な研究が行われているがモデルの記憶容量はこの問題に深く関わっている。[7][18][19] 神経回路網では入力と出力の間の関数を学習する際に、その関数(一般に非常に複雑)そのものを学習するわけではなく、数多くの例を見て、その例に適合するようにシナプス荷重やしきい値を変化させるという帰納的な学習を行っていると考えられる。それでは帰納的学習においてどれだけの例題を学習すれば例題以外の事象に対しても十分に正しい答を出すようにできるかという問題が生じる。見せるべき例題の数は一般に、モデルの記憶容量(正確にはVC次元と呼ばれる量)の関数で書けることかわかっているが、定性的に言えば記憶容

量が增大すればするほど見せるべき例題の数も増大する。

さてここでは2層及び3層の神経回路網の容量について考える。今までに2層の回路については組み合わせ論的な手法を用いて Cover^[8] によって求められている。また Gardner^[10] はスポンガラスにおけるシフト対称性の仮定を用いて解析を行った。3層の回路については最近になって Mitchison ら^[14] が Cover の結果を応用して容量の上限を求めた。まず §2, §3 では Cover と Mitchison の結果を述べ、§4 では中間層の疎符号化と入力パターンของกลุ่ม化によって3層回路の容量のよい下限を達成する回路網を構成する。

§2 2層回路の容量

図3.1は $y = 1(\sum w_i x_i - \theta)$ なる入出力関係をもつ素子である。ただし $x_i \in \mathbb{R}$ とし、 $1(u)$ は

$$1(u) = \begin{cases} 1 & \dots u \geq 0 \\ 0 & \dots u < 0 \end{cases}$$

なるしきい値関数とする。

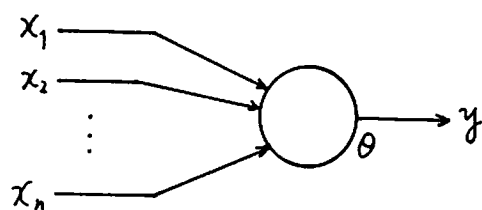


図3.1 しきい素子

この素子に入出力パターン x^1, \dots, x^m と対応する y^1, \dots, y^m を記憶させるとパターン集合 $\{x^m\}$ は 0 を出力するものの集合 X_0 と 1 を出力するものの集合 X_1 とに「2分割」される。その場合の数は 2^m 通り考えられる。

一方図3.1の素子で実現可能な「2分割」というのは $\{x^m\}$ を $n-1$ 次元超平面で分割できるものに限られる。その場合の数は次のとおりである。

定理3.1 (Cover^[8]) \mathbb{R}^n 上に m 個の点が一般の位置に配置されているとする。その点集合を原点を通る $n-1$ 次元超平面で2分割する場合の数は、

$$C(m, n) = 2 \sum_{i=0}^{n-1} \binom{m-1}{i}$$

となる。

証明 帰納法を用いる。 m 個の点が配置されていて超平面で2分割されているとき、新しい1点を付け加えるとどうなるかを考えると次の2つの場合が考えられる。

(a) 超平面を、 m 個の点の2分割を変えないように動かすと、新たにつけ加えられた点を1を出力する側にすることも0を出力する側にすることもできる。(図3.2(a))

(b) 超平面を、 m 個の点の2分割を変えない範囲内で動かしても、新たにつけ加えられた点は1を出力する側か0を出力する側のどちらか一方の側にしか属せない。(図3.2(b))

(a)が起きるのは 超平面が新しい点を通るように動かせる場合であり、その場合の数は新しい点と原点とを

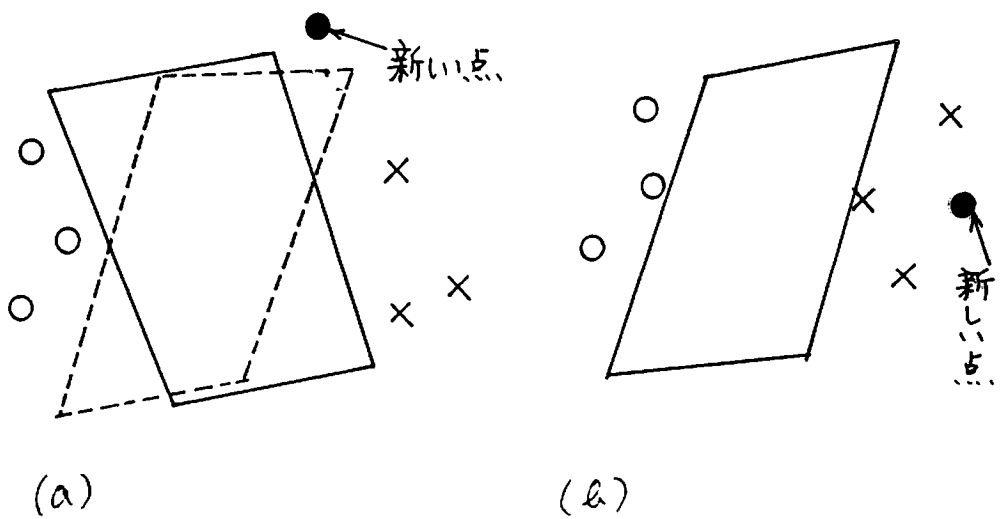


図3.2 新しい点を追加したときの2分割

結ぶ直線に垂直な超平面に m 点を射影したときに, その m 点が線形分離可能となっている場合の数 $C(m, n-1)$ に等しい. 従って漸化式

$$C(m+1, n) = C(m, n) + C(m, n-1)$$

が成り立ち, 初期条件 $C(m, 1) = C(1, n) = 2$ より定理の式を得る. \square

この定理は $C(m, n)$ の値が (点の位置が一般の位置にあるという仮定を除いて) 点の配置によらないことを示しており興味深い.

この定理では超平面が原点を通る, つまり図 3.1 の素子のしきい値が 0 であるという仮定をしたがしきい値が 0 でない場合は次元を 1 つ増やして考えればよい.

この定理を用いると次の定理が成り立つ.

定理 3.2 (Cover) 図 3.1 の入力パターンが点として一般の位置にあり, 各入力パターンに対応する出力パターンが確率 $\frac{1}{2}$ ずつで 0 か 1 をランダムにとるとき, m 個のパターンを識別できる確率は,

$$\lim_{n \rightarrow \infty} P(m, n) = \begin{cases} 1 & \dots \frac{m}{n} < 2 \\ 0 & \dots \frac{m}{n} > 2 \end{cases}$$

となる。

証明 $P(m, n) = \frac{C(m, n)}{2^m}$ となり, m, n が十分大きいときに, この確率分布を正規近似すると, m 個の点から $n = \frac{m}{2} + \frac{\alpha}{2}\sqrt{m}$ 次元で線形分離可能である確率は

$$P\left(m, \frac{m}{2} + \frac{\alpha}{2}\sqrt{m}\right) \sim 1 - F(\alpha)$$

となる。従って $\forall \varepsilon > 0$ に対し,

$$\lim_{n \rightarrow \infty} P(2n(1+\varepsilon), n) = 0$$

となり, 一方

$$\lim_{n \rightarrow \infty} P(2n(1-\varepsilon), n) = 1$$

となる。□

この結果から2層回路の容量が求められる。

系3.3 入力層が n 次元, 出力層が n 次元の2層の神経回路網で、パターンの各成分は確率 $\frac{1}{2}$ ずつで $0, 1$ をランダムに決めるとする。このとき入力パターンが一般の位置にあると仮定すれば、容量は n が十分大きいとき, $2n$ となる。

ただし n は n の多項式オーダーであるとする。

証明 出力層の素子は各々独立に荷重ベクトルを決めることができるから定理 3.2 の結果がそのまま使える。 n を n の多項式オーダーとしたのは、すべての素子について線形分離可能となる確率 $P(m, n)^n$ が n を大きくすると変わってきてしまうためである。 \square

以上の議論は n が大きい場合の漸近的かつ確率的なものである。漸近論を使わないとすると、どのように点を配置しても（ただし一般の位置に配置する）、必ず線形分離可能となるためのパターンの個数の最大値は定理 3.1 より $n+1$ 個（しきい値を 0 に限らない場合）となる。

§3 3層回路の容量の上限

以下では図 3.3 のように入力層 n 次元, 中間層 n 次元, 出力層 l 次元となる 3層回路について考える。この回路にランダムなパターンを記憶させたときの容量の上限は Cover の結果を使って次のような上限をもつ。

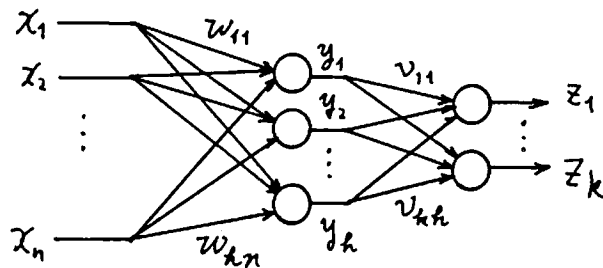


図 3.3 3層の神経回路モデル

定理 3.4 (Mitchison [14]) 入力層が n 次元, 中間層が h 次元, 出力層が k 次元の神経回路網で入力パターンと出力パターンの各成分は確率 $\frac{1}{2}$ ずつで 0, 1 をランダムに決めるとする. このとき容量は n が十分大きいとき, $n \geq h$ ならば $O(nh \log h)$ で上からおさえられる. ここで $t = \frac{h}{k} + 1$ である. 特に $k=1$ のとき $O(nh \log h)$ となる.

証明 $k=1$ の場合を考えると各中間層の素子ごとに線形分離可能な 2 分割は $C(m, n)$ 個あるから回路全体の 2 分割の場合の数は $C(m, n)^h$ で上からおさえられる. ただしこのおさえ方は以下の 2 つの意味で非常に粗いおさえ

方である。

(a) k 枚の超平面が異なる組み合わせであっても 2分割としては同じ場合がある

(b) k 枚の超平面で 2分割を行ったからといってその 2分割が中間層と出力層の間の結合で実現可能とは限らない。

さてこうおさえるとパターンに任意に 0 または 1 をわりあてたとき、その 2分割が回路で実現できる確率は、

$$P = \frac{|C(m, n)|^k}{2^m}$$

で上からおさえられる。従って 3層回路の容量の上限

は $\lim_{n \rightarrow \infty} P = 1$ を満たす最大の m として求められる。

P を Stirling の公式で展開すれば定理の結論を得る。

$k > 1$ の場合にも同様に超平面による分割の総数を $|C(m, n)|^k |C(m, k)|^k$ で上からおさえ、さらに $n \geq k$ としてこれを $|C(m, n)|^{k+k}$ で上からおさえることにより、定理の結論を得る。□

§4 3層回路の容量の下限

(1) 中間層を疎符号化したモデル

定理3.4の証明の中で述べたように, Mitchison らの与えた上限というのは非常に粗いことがわかる. 一方3層回路の上限としては $O(nh)$ であろうという指摘もある.

§4では n, h が十分大きいとき $O\left(\frac{nh}{\log(nh^2)}\right)$ の容量をもつ回路網を実際に構成することで3層回路の容量の下限を与える. 基本的な考え方はスプース化によって容量を増やすということであるが, 今考えている問題では入力層と出力層はスプースではないから中間層をスプースにするしかない. しかしこうすることによって入力層と中間層の間の容量を増やすことはできるが中間層と出力層の間の容量は変わらない. そこで中間層と出力層の間は少ないパターン数を記憶させ, 入力パターンをグループ分けし, そのグループと中間層の1つのパターンとの間で連想記憶を行う. ただし1つのグループに属するパターンの数をそれほど多くできないことにも注意する必要がある.

ここでモデルについてさらに詳しく述べることにする。 m 個の入カパターンは α -スパースでランダムに与えられるとする。 3層回路の容量を考えると $\alpha = \frac{1}{2}$ である。 ここで l 個の中間層のパターン y^1, \dots, y^l を α -スパースでランダムに決める。 そして入カパターンを同じ出カパターンをもつもの同士 $M = \frac{m}{l}$ 個ずつ l 個のグループ X_1, \dots, X_l に分け、グループ X_{α} に属するパターンを入力すると中間層に y^{α} があらわれ、それに対して出力 z^{α} が出力されるようにする。(図 3.4) ここで l としては出カパターンの個数以上である必要があるかこれについてはあとで考えることにする。

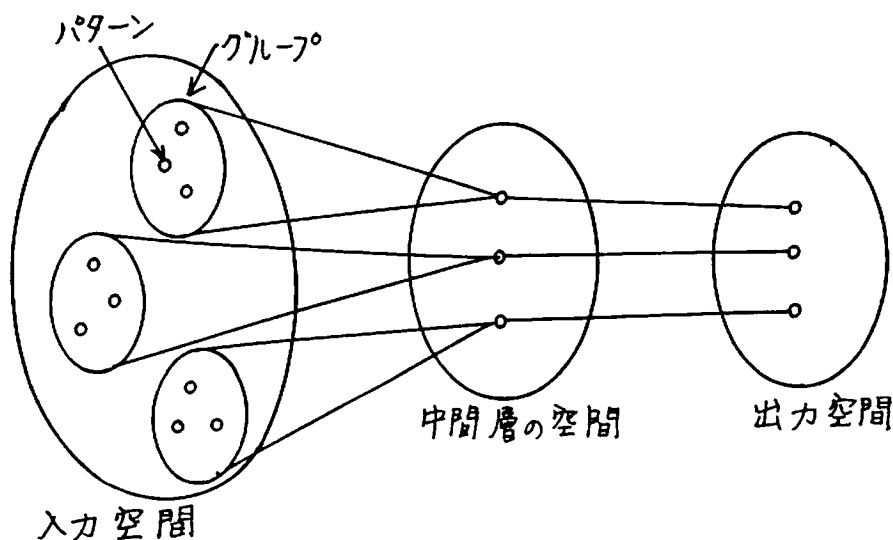


図 3.4 パターンの写像

また x の添え字を j に変えて、グループ X^α に属するパターンを $x^{\alpha,1}, x^{\alpha,2}, \dots, x^{\alpha,M}$ と書くことにする。

入出力関係は、入力層から中間層への結合については相関を用いた連想記憶で行ったのと同様に、

$$y_i = f\left(\sum_j w_{ij} x_j - \theta\right)$$

とし、シナプス荷重は

$$w_{ij} = \frac{1}{n} \sum_{\alpha=1}^l \sum_{\nu=1}^M y_i^\alpha x_j^{\alpha,\nu}$$

ととることにする。中間層から出力層への結合については

$$z_k = f\left(\sum_i v_{ik} y_i - \hat{\theta}\right)$$

であるとする。シナプス荷重については後から考える。

(2) モデルの容量

中間層のパターンの数 l とスパース度 h を固定したときの容量は次のようになる。

定理 3.5 中間層のパターンが l 個で h -スパースのときの容量は、 n が十分大きいとき、

$$C = \frac{n}{8 \left\{ \frac{1}{l} + h(1-h) \right\} \log(hC)}$$

の解で与えられる。

証明 証明の手法としては定理 2.1 や定理 2.4 と同様にやればよい。グループ X_1 の第 1 要素 $x^{1,1}$ からパターン y^1 が想起できるのは $u_i = \mathbb{V} x^{1,1}$ として、 $y_i^1 = f(u_i - \theta)$ となる場合である。以下入力パターンを a -スパースであるとする。

$$\begin{aligned} u_i &= \frac{1}{n} \sum_{\alpha=1}^L \sum_{\nu=1}^M y_i^\alpha x^{\alpha,\nu} x^{1,1} \\ &= \frac{1}{n} y_i^1 x^{1,1} \cdot x^{1,1} + \frac{1}{n} y_i^1 \sum_{\nu \neq 1} x^{1,\nu} x^{1,1} \\ &\quad + \frac{1}{n} \sum_{\alpha \neq 1} y_i^\alpha \sum_{\nu} x^{\alpha,\nu} x^{1,1} \\ &= \{ a(1-a) + L_i \} y_i^1 + N_i. \end{aligned}$$

ただし、

$$\begin{aligned} L_i &= \frac{1}{n} \sum_{\nu \neq 1} x^{1,\nu} x^{1,1}, \\ N_i &= \frac{1}{n} \sum_{\alpha \neq 1} y_i^\alpha \sum_{\nu} x^{\alpha,\nu} x^{1,1}. \end{aligned}$$

ここで a をスパースでないとして仮定すれば中心極限定理が成立し、 L_i, N_i は漸近的に正規分布となる。平均はともに 0 であり、分散は

$$\begin{aligned} \mathbb{V}[L_i] &= \frac{M}{n} a^2 (1-a)^2 = \sigma_L^2 \\ \mathbb{V}[N_i] &= \frac{M \cdot L}{n} b(1-b) a^2 (1-a)^2 = \sigma_N^2 \end{aligned} \quad (3.1)$$

となる。簡単のためしきい値を $-h a(1-a) + \frac{1}{2} a(1-a)$

とすると、 y_i^1 の条件付誤り確率は、

$$Pr(y_i^1 \neq f(u_i - \theta) | y_i^1 = 1-h) = F\left(\frac{a(1-a)}{2\sigma}\right)$$

$$Pr(y_i^1 \neq f(u_i - \theta) | y_i^1 = -h) = F\left(\frac{a(1-a)}{2\sigma_N}\right) \\ < F\left(\frac{a(1-a)}{2\sigma}\right)$$

となる。従ってすべてのパターンについての誤った成分の個数の期待値は

$$E = m h F\left(\frac{a(1-a)}{2\sigma}\right)$$

でおさえられる。従って $E \rightarrow 0$ となるのは (2.8) 式の近似 ($n \rightarrow \infty$)

を使って、

$$\frac{a^2(1-a)^2}{8\sigma^2} > \log(mh)$$

が成り立つことが十分。(3.1) の値を代入すれば、

$$\frac{n}{8\left\{1+h(1-h)\right\}M} > \log(mh)$$

$$\therefore m < \frac{n}{8\left\{1+\frac{1}{2}+h(1-h)\right\}\log mh}$$

が成り立つ。 \square

この証明では λ の値の与え方において計算の簡単のための仮定をおき、誤り確率も上からおさえたので、実際の容量はさらに多くなる。これらのことを考慮に入れても定数倍増加するだけで、オーダーは変わらない。

定理 3.5 は λ が大きくなればなるほど、また n が小さくなればなるほど容量が大きいことを示している。

次に中間層から出力層への想起を考える。Cover の結果を使えば、パターンベクトルが一般の位置にあれば λ としては $2n$ までとることができる。実験的にはほとんどの場合に一般の位置にあると見てよいが、次のようなパターンベクトルを並べた行列を考えるとある条件で正則になることが言えて、その後示されるように λ としては n 以上にはできることがわかる。

定理 3.5 より中間層のパターンはスパースであるほどよいから $\frac{c}{n}$ -スパースであるとする。

$$y_i^\alpha = \begin{cases} 1 & i \equiv \alpha, \alpha+1, \dots, \alpha+c-1 \pmod{n} \\ 0 & \text{それ以外} \end{cases}$$

とくと次の補題が成り立つ。

補題 3.6 行列 $Y = (y_{ij}^x)$ が正則であるための必要十分条件は $c=1$ かまたは n と c が互いに素となることである。

証明

$$T = \begin{bmatrix} 0 & & & & 1 \\ 1 & & & & \\ & \ddots & & & \\ 0 & & \ddots & & \\ & & & 1 & 0 \end{bmatrix}$$

とおくと,

$$Y = f(T) = I + T + T^2 + \dots + T^{c-1}$$

と書ける。 T の固有方程式は $\varphi(t) = t^n - 1$ であり、固有値は、 $e^{2\pi i \frac{j}{n}}$ ($j=0, \dots, n-1$) となるから Y の固有値は $f(e^{2\pi i \frac{j}{n}})$ ($j=0, \dots, n-1$) となる。 Y が正則であるための必要十分条件は Y が固有値 0 をもたないことである。

$j=0$ のとき $f(1) \neq 0$ だから $t \neq 1$ を仮定すると,

$$f(t) \neq 0 \Leftrightarrow (1-t)f(t) \neq 0 \Leftrightarrow 1-t^c \neq 0.$$

従って $j=1, \dots, n-1$ に対して

$$1 - e^{2\pi i \frac{j}{n} c} \neq 0$$

であればよい。これは c と n が互いに素のとき、またそのときに

限って成り立つ。 \square

Y が正則ならば、中間層と出力層の間のシナプス荷重が存在する。つまり出力パターン z_i^α が 0, 1 の 2 値をとるとき、

$$z_i^\alpha = v_i \cdot y^\alpha \quad (\alpha = 1, \dots, k)$$

なる k 次元ベクトル v_i が存在する。これを利用すれば、

$$z_i^\alpha = 1 \left(v_i \cdot y^\alpha - \frac{1}{2} \right)$$

と書ける。

定理 5 の結果より C の値は「せいほど」よいから、 $Y=I$ (I は単位ベクトル) のときが容量最大であり、中間層のパターンとしては k 個までとれる。これを定理 3.5 の結果に代入することによって、次の定理を得る。

定理 3.7 入力層 n 次元、中間層 k 次元、出力層 k 次元の 3 層回路網にランダムなパターンを記憶させたときの容量の下限は次式のようになる。

$$C = \frac{nk}{16 \log(nk^2)}.$$

証明 略 \square

さて以上の議論では入力パターンを M 個ずつ k 個のグループに分けたから出力パターンは k 以下でなければならぬ。

定理3.7では $l=n$ としたから、出力層が長次元のとき、 $n \geq 2^k$ を満たす必要がある。

ここで述べたモデルは非常に単純であり、実際の子層回路の容量はこれより多くなるだろうと思われる。まず第1に中間層のパターンはスパースであるという仮定を除いて全くランダムにとってよかった。第2には出力素子の数が非常に少ない場合（例えば $l=1$ のとき）、同じ出力を出す入力パターンの集合でも、異なるいくつかのグループに分けることが必要となるが、このモデルでは全くランダムにとってよいとしていることである。第3に、入力層から中間層の間で相関学習の連想記憶を行っている。容量を増加させるような他の結合が存在するかもしれない。

第2点に関して、 l が大きい場合 ($\log_2 n$ 程度) を考えると、このモデルは非常によい下限を与えていると思われる。実際 $k = \log_2 n$ のとき、Mitchisonらの与えた上限の値は $O(nk)$ となる。実際の容量はこれより少ないはずだから定理3.7で与えた容量は実際の容量に近いと思われる。

このことは長が大きくなるとグループ分けのしかたに自由度がなくなることが1つの原因かもしれない。従って長=1の場合にはグループ分けをうまくやって容量を増やすことができる可能性もある。

(3) モデルの意味

ここではこのモデルが情報処理システムとして持つ意味について考えてみる。画像や言葉といった個々の「事物」または「事象」がシステムに入力され、それに対する「行動」または「反応」を出力するシステムを考える。これを2層の回路で実現しようとする、 $2n$ 個(n は入力ベクトルの次元)の「事物」に対する「行動」しか記憶できないことになる。一方3層回路にして、中間層の各素子に「事物」の「種類」や(抽象)概念のようなものを割り当て、「事物」と「種類」との関係を記憶し、「種類」と「行動」との間は独立して学習を行うことにする。すると「種類」の数としては中間層の素子の数だけ記憶でき、1つの「種類」に対して $O\left(\frac{n}{\log(nk^2)}\right)$ 個の「事物」を対応させられ、

全体として $O\left(\frac{nr}{\log(nr^2)}\right)$ 個の「事物」を記憶することからできることになる。またこの場合、同じ「種類」に属する「事物」同士は全く相関を持っていないとしてもよい。

このモデルは従来から言われている「おぼあさ細胞」、つまり個々の「事物」に対して1つの細胞が割り当てられるとするモデルとは似て非なるものであることがわかる。

第4章 競合機構をもつ連想記憶モデル

§1 基本競合系

第2章で考えたスパースな連想記憶モデルのうち入力もスパースであるものを考える。これをフィードバックのない2層回路で実現しようとする。2つのパターンが重なりあって入力されると重なりあったパターンを出力するという性質をもっている。それを起こらないようにするには入力層あるいは出力層に競合機構を設けて素子間に競合を起こさせ出力の数をスパースに保つことが必要となる。

そこでまずもととなる基本競合系^{[1][2]}の性質を調べ

る。(図4.1)

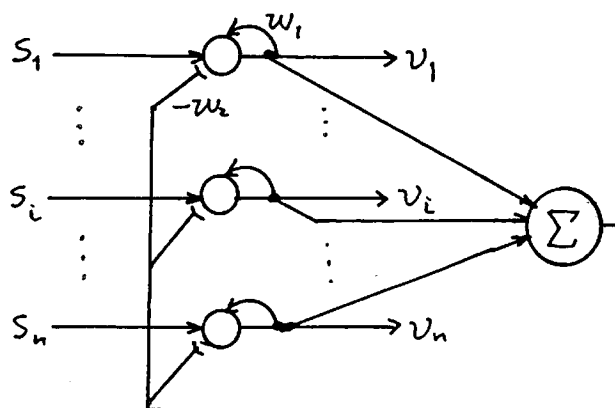


図4.1 基本競合系

各素子は内部状態 u_i をもち、 $v_i = 1(u_i)$ によって v_i を出力する。また自分自身に重み w_1 でフィードバックをかけ、(自身を含む) 各素子に重み $-w_2$ で抑制をかけるような回路網である。系の方程式は

$$\dot{u}_i = -u_i + w_1 v_i - w_2 \sum_j v_j + s_i \quad (4.1)$$

とあり、入力 s_i は一定で

$$s_{\min} \leq s_i \leq s_{\max}$$

を満たすとする。

出力パターンがスパースとなるためにはダイナミクスの結果 s_1, \dots, s_n のうち大きいものから k 個の入力が入る素子だけが興奮するように w_1, w_2 の値を決める必要がある。

($\frac{k}{n}$ -スパース)

以下の解析で l をある時刻で発火している素子の数であるとし、 $l > k$, $l = k$, $l < k$ の各場合について (4.1) 式の振る舞いを調べ $l = k$ が平衡状態になるための条件を求め、一般にはそれができないことを示す。

まず $l > k$ の場合、発火している素子について (4.1) 式は

$$\dot{u}_i = -u_i + w_1 - w_2 l + S_i < w_1 - w_2 (l+1) + S_{max}$$

だから,

$$w_1 < w_2 (l+1) - S_{max} \quad (4.2)$$

としてやれば $l > l$ は平衡状態ではあり得ない。

次に $l = l$ の場合, 発火している素子について (4.1) 式は

$$\dot{u}_i = -u_i + w_1 - w_2 l + S_i > -u_i + w_1 - w_2 l + S_{min}$$

だから

$$w_1 > w_2 l - S_{min} \quad (4.3)$$

としてやれば u_i が正から 0 に向かって減少していても,

$u_i = w_1 - w_2 l + S_i > 0$ で平衡状態になるため u_i は負にならない。

また $l = l$ の場合, 発火していない素子について (4.1) 式は

$$\dot{u}_i = -u_i - w_2 l + S_i < -u_i - w_2 l + S_{max}$$

だから,

$$w_2 l > S_{max} \quad (4.4)$$

としてやれば u_i が負から 0 に向かって増加していても,

$u_i = -w_2 l + S_i < 0$ で平衡状態になるため u_i は正に

ならない。

次に $l < l_c$ のときを考えると、発火していない細胞について、

$$\dot{u}_i = -u_i - w_2 l + s_{min} > -w_2 (l_c - 1) + s_{min}$$

だから、

$$w_2 (l_c - 1) < s_{min} \quad (4.5)$$

としてやれば $l < l_c$ は平衡状態ではあり得ない。

以上のことからダイナミクスの結果として $l = l_c$ となるのは、(4.2) ~ (4.5) 式の条件を満たす必要がある。逆に初期状態において u_i がすべて同じ初期値をとるものとするとき、(4.1) 式の形から \dot{u}_i の大きさの順序は s_i の大きさの順序と一致するため、 s_i の値が大きいものほどはげしく増加することになる。従って (4.2) ~ (4.5) 式を満たすように w_1, w_2 を決めれば s_1, \dots, s_n のうちの大きいものから l_c 個の入力が入る素子が発火し、それ以外の素子は発火しないことが言える。

さてそれでは (4.2) ~ (4.5) 式の条件を満たす w_1, w_2 が常に存在するかどうかについて考えてみる。条件をグラフで表すと

図4.2 のようになり, (4.2) ~ (4.4) 式の条件を満たす w_1, w_2 は必ず存在するが, (4.5) の条件は, 次のとき満たされない。

$$\frac{S_{\min}}{k-1} < \max\left(\frac{S_{\max}}{k}, S_{\max} - S_{\min}\right).$$

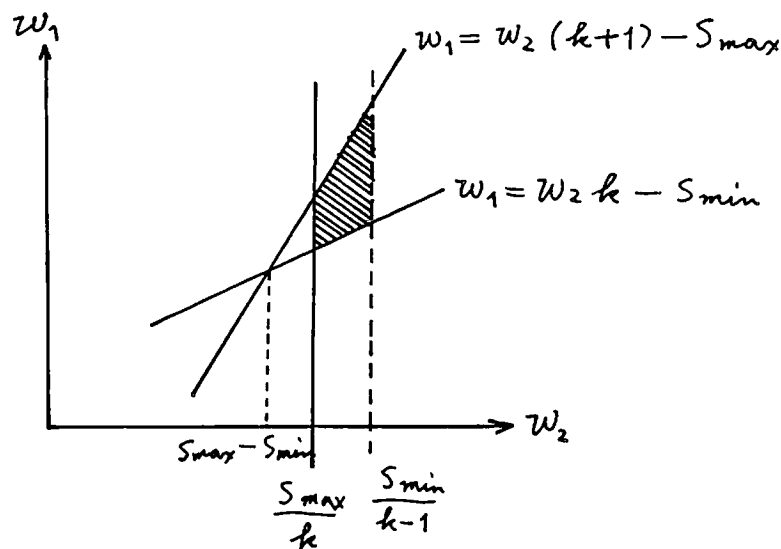


図4.2 基本競争系のパラメータの範囲

(4.2)式 ~ (4.4)式の条件を満たす w_2 をとったとき,

$$k'w_2 > S_{\min}$$

を満たす最小の k' をとると, 系は $k' \sim k$ の範囲で平衡状態となる。この場合も u_i がすべて正という初期値から始めれば $l = k$ が平衡状態となる。

一般には $l = k$ だけが平衡状態となるような w_1, w_2 が

存在しないことが示された。つまりいくつかの素子が互いに拮抗しているときには、「すくみ」の状態に陥りやすいことを表している。

ただし $k=1$ の場合だけは特別で $S_{min} > 0$ であれば (4.5) 式の条件は必ず満たされることになる。

以上のことを定理の形でまとめる。

定理 4.1 基本競合系は w_1, w_2 が次の方程式を満たすとき、 $k' \sim k$ 個の発火数を平衡状態としてもつ。

$$w_2 k - S_{min} < w_1 < w_2 (k+1) - S_{max}$$

$$\frac{S_{max}}{k} < w_2 < \frac{S_{min}}{k'-1}$$

特に u_i が正の定数で初期化されているとき、入力のうち大きいものから k 個が入ってくる素子だけが興奮することになる。

証明 略 \square

§2 競合機構と連想記憶モデルの疎符号化

ここでは §1 で述べた競合機構とスパースな連想記憶モデルを組み合わせた図 4.3 の系の振る舞いを調べる。

連想記憶モデルとしては特に自己想起型、つまり X^m または

それに近いパターンを入力として x^{μ} を出力させるモデルを考える。

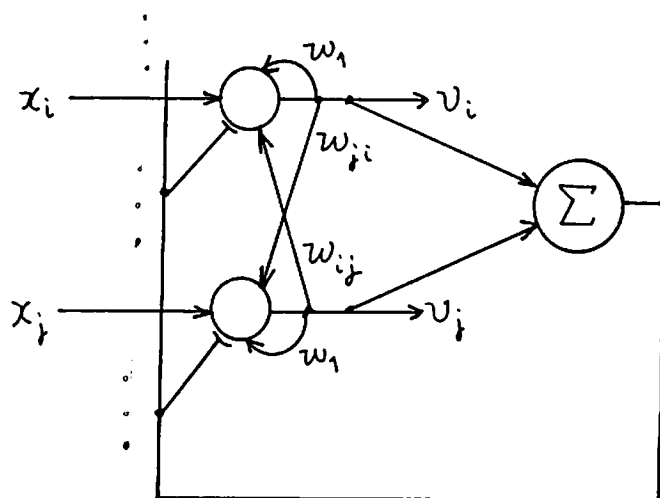


図4.3 競合機構と連想記憶を組み合わせたモデル

また、パターンは α -スパースであるとするか、 α としては非常にスパースなもの ($O(\frac{\log n}{n})$) を考えることにし、0-1モデルとする。図4.3では基本競合系の結合の他に相関を用いた連想記憶の結合 $w_{ij} = \frac{\alpha}{n} \sum_{\mu} x_i^{\mu} x_j^{\mu}$ があるものとする。この系にはまず外部から入力 x が入力され、内部状態 u_i は $x_i = 1$ のとき正の定数に、 $x_i = 0$ のとき負の定数に初期化される。その後は入力は途端えて系の内部結合だけでダイナミクスが進行するとする。つまり

$$u_i = -u_i + w_1 v_i - w_2 \sum_j v_j + \frac{\alpha}{n} \sum_{\mu, j \neq i} x_i^{\mu} x_j^{\mu} v_j \quad (4.6)$$

に従うものとする。このモデルは定性的には次のように働く。

パターン数が少ないときには主として基本競合系として、パターン数が多くなるにつれ、連想記憶として働く。

それでは系がうまく働くように w_1, w_2 の値を決めることにする。

パターンの成分のうち 1 をとるものの数を $l = n\alpha$ とおき、ある時刻で発火している素子の数を l とする。

まずパターン x^1 を入力したときそれが平衡状態となるための条件を考える。 $x_i^1 = 1$ なる i について (4.6) 式は

$$\dot{u}_i = -u_i + w_1 - w_2 l + \alpha (\alpha + N_i),$$

となる。ただし $N_i = \frac{1}{n} \sum_{\substack{m=1 \\ m \neq i}}^n x_i^m x_j^m x_j^1$ とし、

$$N_{\min} \leq N_i \leq N_{\max}$$

を満たすとする。ここで

$$w_1 - w_2 l + \alpha (\alpha + N_{\min}) > 0 \quad (4.7)$$

であれば u_i は負にならない。また $x_i^1 = 0$ なる i については

(4.6) 式は

$$\dot{u}_i = -u_i - w_2 l + \alpha N_i$$

となり、

$$w_2 l > \alpha N_{max} \quad (4.8)$$

であれば u_i は正にならない。

従って (4.7) 式, (4.8) 式を満足せばパターン x^1 を入力したときにそれが平衡状態となる。

次には記憶されているパターンと少し異なるベクトルを入力したときの振舞いを調べる。まず現在発火しているか $x_i^1 = 0$ であるような素子について (4.6) 式は

$$\dot{u}_i = -u_i + w_1 - w_2 l + \alpha N_i < w_1 - w_2 l + \alpha N_{max}$$

だから,

$$w_1 - w_2 l + \alpha N_{max} < 0 \quad (4.9)$$

を l の付近のできるだけ多くの l に関して成り立たせてやればその範囲でこの素子は $u_i > 0$ を平衡状態として持たない。

一方現在発火していないか $x_i^1 = 1$ である素子について (4.6) 式は

$$\dot{u}_i = -u_i - w_2 l + \alpha \{ a(1-p) + N_i \}$$

(ただし p は $x_i^1 = 1$ となる成分のうち $v_i^1 = 1$ となっているものの割合)

となるから,

$$w_2 l < \alpha \{ a(1-p) + N_{min} \} \quad (4.10)$$

を l の付近でできるだけ多くの l に関して成り立たせてやればその範囲でこの素子は $u_i < 0$ を平衡状態としてもたない。

以下では (4.9) 式や (4.10) 式が満たされるような l の範囲をおおまかに調べる。

まず計算を具体的にを行うために $a = \frac{\log n}{n}$ とする。

N_{\max} や N_{\min} を評価するために N_i の分布を考えると、定理 2.1 の証明の途中結果を使うと、 m が記憶容量の付近で、

$$E[N_i] = ma^3 = O\left(\frac{\log n}{n}\right) = O(a),$$

$$\sqrt{V[N_i]} = \sqrt{\frac{ma^3}{n}} = O\left(\frac{\sqrt{\log n}}{n}\right) < O(a).$$

(ここで N_i の平均が 0 でないのは 0-1 モデルを考えているからである。)

従って容量の付近では

$$N_{\max} \simeq N_{\min} \simeq \beta a$$

とみなせる。

一方 w_1, w_2 としては (4.7) 式, (4.8) 式の不等式をほぼ等式で満たすようにとる。つまり

$$w_2 = \frac{\alpha N_{\max}}{k} \simeq \frac{\alpha \beta}{k} a, \quad (4.11)$$

$$w_1 = \alpha (N_{\max} - N_{\min} - a) \sim -\alpha a \quad (4.12)$$

とする。これを(4.9)式に代入すると、

$$\begin{aligned} l &> \left(1 - \frac{a + N_{\min} - N_{\max}}{N_{\max}}\right) l \\ &\sim \left(1 - \frac{1}{\beta}\right) l \end{aligned} \quad (4.13)$$

の範囲で0を入力すべきところを1と入力(してしまつた誤り)を修正する方向に進む。(4.11), (4.12)式を(4.10)式に代入すると、

$$\begin{aligned} l &< \frac{a(1-p) + N_{\min}}{N_{\max}} l \\ &\sim \left(1 + \frac{1}{\beta}(1-p)\right) l \end{aligned} \quad (4.14)$$

の範囲で1を入力すべきところを0と入力(してしまつた誤り)を修正する方向に進む。

(4.13)式と(4.14)式はともに β を増やす、つまり記憶パターンを増やすと誤りを修正できる範囲が狭くなっていくことを表し、 β が十分小さければ“ l のまわりの広い範囲で誤りを修正できることを表している。

また(4.12)式は一般に w_1 が負になる、つまりこのモデルでは自己フィードバックが自己抑制として働くことを意味しており、

No. 53

Date

興味深い。

第5章 結論

本論文では スパースな情報表現をもつ神経回路モデルの特性を調べた。

まず第1に2層の神経回路網で出力パターンをスパース化しておくことにより記憶容量を増やすことができることを示し、この性質が入力パターンがスパースであるかどうかには依存しない性質であることを示した。また引き込み領域の広さによって容量がどう変化するかを調べ、一般に1を入力すべき成分を0と間違えるような誤りを訂正するには容量を大きく減らす必要があることを示した。

第2にはスパース化によって容量が増えることを利用し、入力と出力がスパースでない層回路網の容量の下限を与えるモデルを提案した。

第3にはスパース化を実現するための基本競合系の特性を調べ、基本競合系と連想記憶を組み合わせたモデルについて考察した。

今後はこのようなモデルを複合的に組み合わせ^[15]て海馬な

どの複雑なシステムのモデル化およびその解析を行って
脳のメカニズムを解明するとともに、それらの結果を逆に新しい
コンピュータアーキテクチャに応用することを考えた。本論文
では具体的な脳から少し離れた視点で教理工学的な
解析を行ったが、脳のモデル化などを行う際には、(枝
葉末節でない本質的な)生理学的な実験事実や、(十
分科学的な)認知科学的視点を考慮していく必要が
あるだろう。

謝 辞

本論文を書くにあたり指導教官である甘利俊一教授には御多忙にもかかわらず数々の御助言・御指導を頂き、ここに深く感謝の意を表します。また、いろいろな議論に参加して下さいた倉田耕治助手をはじめとする第1講座研究室の皆様にも感謝致します。

参考文献

- [1] 甘利 俊一 (1978): 神経回路網の数理. 産業図書.
- [2] Amari, S. and Arbib, M.A. (1977): Competition and cooperation in neural nets. In Metzler, J. (Ed.), *Systems neuroscience*. New York: Academic Press.
- [3] Amari, S. and Maginu, K. (1988): Statistical neurodynamics of associative memory. *Neural Networks*, Vol. 1, No. 1, pp. 63–73.
- [4] Amari, S. (1989): Characteristics of sparsely encoded associative memory. *Neural Networks*, Vol. 2, No. 6, pp. 451–457.
- [5] Amit, D.J., Gutfreud, H. and Sompolinsky, H. (1985): Spin-glass models of neural networks. *Phys. Rev. A*, Vol. 2, pp. 1007–1018.
- [6] Anderson, J.A. (1972): A simple neural network generating interactive memory. *Mathematical Biosciences*, Vol. 14, pp. 197–220.
- [7] Baum, E.B. and Haussler, D. (1989): What size net gives valid generalization? *Neural Computation*, Vol. 1, pp. 151–160.
- [8] Cover, T.M. (1965): Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition. *IEEE Trans. Electron. Comput.*, Vol. 14, pp. 326–334.
- [9] Gardner, E. (1986): Structure of metastable states in the Hopfield model. *J. Phys. A: Math. Gen.*, Vol. 19, pp. L1047–L1052.
- [10] Gardner, E. (1988): The space of interactions in neural network models. *J. Phys. A: Math. Gen.*, Vol. 21, pp. 257–270.
- [11] Hopfield, J.J. (1982): Neural networks and physical systems with emergent collective computational abilities. *Proc. National Academy of Sciences, USA*, Vol. 79, pp. 2445–2558.
- [12] Kohonen, T. (1972): Correlation matrix memories. *IEEE Trans. Comput.*, Vol. 21, pp. 353–359.
- [13] McEliece, R.J., Posner, E.C., Rodemich, E.R. and Venkatesh, S.S. (1987): The capacity of the Hopfield associative memory. *IEEE Trans.*

- Inf. Theory*, Vol. 33, pp. 461–482.
- [14] Mitchison, G.J. and Durbin, R.M. (1989): Bound on the learning capacity of some multi-layer networks. *Biol. Cybern.*, Vol. 36, pp. 646–658.
- [15] Miyashita, Y. and Chang, H.S. (1988): Neuronal correlate of pictorial short-term memory in the primate temporal cortex. *Nature*, Vol. 331, pp. 68–70.
- [16] Nakano, K. (1972): Associatron — a model of associative memory. *IEEE Trans. Sys. Man Cybern.*, Vol. 2, pp. 381–388.
- [17] Palm, G. (1981): On the storage capacity of an associative memory with randomly distributed storage elements. *Biol. Cybern.*, Vol. 36, pp. 646–658.
- [18] Vapnik, V.N. and Chervonenkis, A.Ya. (1971): On the uniform convergence of relative frequencies of events to their probabilities. *Theory Prob. Appl.*, Vol. 16, pp. 264–280.
- [19] Valiant, L.G. (1984): A theory of learnable. *Comm. ACM*, Vol. 27, No. 11, pp. 1134–1142.
- [20] Weisbuch, G. (1985): Scaling laws for the attractors of Hopfield networks. *J. Phys. Lett.*, Vol. 46, pp. 623–630.
- [21] Willshaw, D.J. and Longuet-Higgins H.C. (1970): Associative memory models. In Meltzer, B. and Michie, O. (Eds.), *Machine Intelligence*, Vol. 5. Edinburgh: Edinburgh University Press.