



HOKKAIDO UNIVERSITY

Title	2004年度 混沌系工学特論講義ノート
Author(s)	井上, 純一; Inoue, Jun-ichi
Description	当講義資料は著者のホームページ http://chaosweb.complex.eng.hokudai.ac.jp/~j_inoue/ からダウンロードできます。 http://chaosweb.complex.eng.hokudai.ac.jp/~j_inoue/
Issue Date	2004
Doc URL	https://hdl.handle.net/2115/370
Rights(URL)	https://creativecommons.org/licenses/by-nc-sa/2.1/jp/
Type	learning object
File Information	konton2004_3.pdf, 第3回講義ノート



混沌系工学特論 配布資料 #3

担当：井上 純一 (情報科学研究科棟 8-13)

URL : http://chaosweb.complex.eng.hokudai.ac.jp/~j_inoue/

平成 16 年 11 月 8 日

目次

1.11 確率的動作をする神経回路網の連想記憶	26
1.11.1 第 2 のノイズ：熱雑音の効果	26
1.11.2 ニューロンの動作は不確かである	27
1.11.3 平均場近似という考え方	28
1.11.4 局所的重なりの数値解法：平均場アニーリング	28
1.11.5 $p = \mathcal{O}(1)$ の場合の解析解とその相転移	31
1.11.6 $p = \mathcal{O}(N)$ の場合の解析解とその相転移	33

1.11 確率的動作をする神経回路網の連想記憶

1.11.1 第 2 のノイズ：熱雑音の効果

前節までで、パーセプトロンと呼ばれる神経素子を Hebb 則に基づいて多数つなげた場合、その回路網が連想記憶として動作することを学んだ。多数の (神経素子数に比例するオーダーの) パターンを回路網に記憶させた場合、任意のパターンを想起することが出来るか否かは、想起パターン ξ^1 に関する素子 i の内部ポテンシャル：

$$h_i = \frac{1}{N} \sum_{j=1}^N \sum_{\mu=1}^p \xi_i^\mu \xi_j^\mu \xi_j^1 = \xi_i^1 + \frac{1}{N} \sum_{j=1}^N \sum_{\mu \neq 1} \xi_i^\mu \xi_j^\mu \xi_j^1 \quad (62)$$

によって決まり、特に、パターン数 p が $\mathcal{O}(N)$ の場合に上式 (62) 第 2 項は $\mathcal{O}(1)$ の確率変数となるので、この変数の統計的性質を調べることにより、記憶させることのできるパターン数についての情報を得ることができた。この第 2 項はクロストークと呼ばれ、多数のパターンを記憶させてシステムを構成することによって生じ、望ましい機能の発現 (記憶の想起) を妨害するという意味においてノイズである。しかし、このようなシステム構成時に内在してしまうようなノイズとは別個のノイズ — 外部ノイズ — も脳神経系には存在する。例えば、1 つひとつのニューロンの動作は外部からの雑音 (例えば熱的な揺らぎ) によって、必ずしも前に見たような正確な動作を示さない場合があり得る。そこで、ここからはこの種の外部ノイズが連想記憶に及ぼす影響を考えて行くことにしよう。ただし、以下では対称結合の場合に限って話を進める。

1.11.2 ニューロンの動作は不確かである

第 1.1 節で述べたように、神経回路を構成するニューロンの動作は決して信頼のおけるものではない。前節で考えた神経回路網は、 i 番目の神経素子に入力される信号の総和 (内部ポテンシャル) :

$$h_i \equiv \frac{1}{N} \sum_{j \neq i} \sum_{\mu} \xi_i^{\mu} \xi_j^{\mu} S_j \tag{63}$$

が正ならば $S_i = 1$, 負ならば $S_i = -1$ として i 番目の素子の出力が決定された。しかし、神経素子の入出力がこのような決定論的ルールに従って決まるというよりも、むしろ不安定な素子として、その出力を確率的に決定するものとして考えた方が、より現実的な神経素子の振舞いを説明し、その素子を多数つなげた神経回路網の性質を議論する場合にも望ましいのではないかと考えられる。この「確率的」という動作をどのように導入するかは様々あるが、ここでは i 番目の神経素子が $S_i = 1$ の出力をする確率 $P(S_i = 1)$ が

$$P(S_i = 1) = \frac{e^{\beta h_i}}{e^{\beta h_i} + e^{-\beta h_i}} \tag{64}$$

で与えられるような神経素子を考える。ここで、 β は誤作動の程度を表すパラメータであり、 $\beta \rightarrow \infty$ の極限を考えると、 $h_i > 0$ のとき $P(S_i = 1) = 1$, $h_i < 0$ のとき $P(S_i = 1) = 0$ となるから前節までの「決定論的素子」に戻る。逆に $\beta = 0$ の極限では $P(S_i = 1) = 1/2$ となり、「全く信頼できない素子」となってしまう (図 19 参照)。ここまで素子の動作を定義したところで、この確率的動作の「起源」は一体何なのかが知

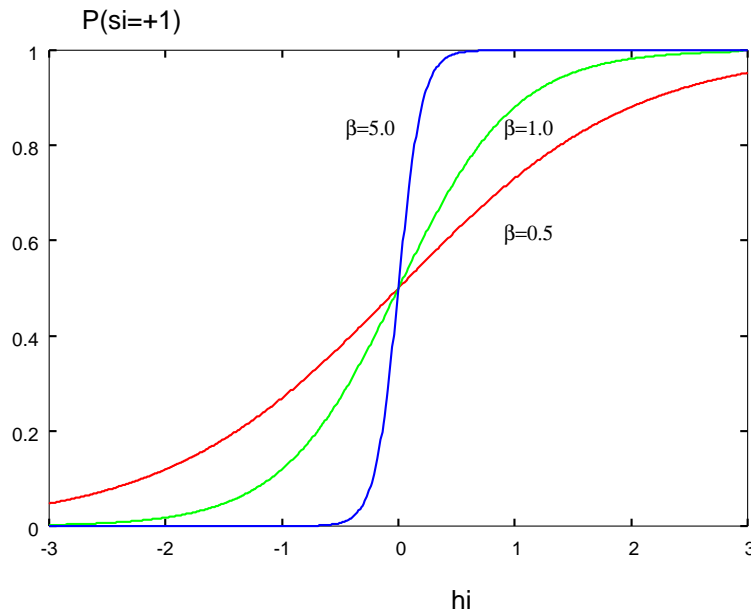


図 19: 内部ポテンシャル h_i の関数としての確率 $P(S_i = 1)$.

りたくなるが、ここではその詳細までは踏み込まないことにする。この種のノイズは [熱的なもの] であるかもしれないし、何らかの [化学的反応によるもの] かもしれない。しかし、ここではそれらの起源が何であれ、そのノイズを全て β の中に押し込めて議論する。ただし、ここで導入した β で表されるノイズが前節までに見たクロストークとは本質的に異なることだけは明らかであろう。

また、このトピックス [連想記憶の数理] の初めに神経素子とノイマン型コンピュータとの比較を行ったが、その際に脳を計算機としてみた場合、「神経回路網は個々の素子の不確かさを膨大な数の素子を結合さ

せることによって補っている」ということを述べた。従って、個々の素子の不確かさを β で導入し、その不確かさを徐々に増加して行った場合、連想記憶という多数の素子どうしの協力に基づく機能がどこまで損なわれずに残るのかを調べることは大変に有意義であろうし、以下では焦点をそこに絞って議論したい。

さて、上記のように素子が確率的動作を示すとするならば、素子 i が $S_i = -1$ をとる確率は

$$P(S_i = -1) = 1 - P(S_i = 1) = \frac{e^{-\beta h_i}}{e^{\beta h_i} + e^{-\beta h_i}} \quad (65)$$

で与えられる。

神経素子が、この確率で動作すると仮定すると i 番目の神経素子の出力値の平均 $\langle S_i \rangle$ は

$$\langle S_i \rangle = (+1) \times \frac{e^{\beta h_i}}{e^{\beta h_i} + e^{-\beta h_i}} + (-1) \times \frac{e^{-\beta h_i}}{e^{\beta h_i} + e^{-\beta h_i}} = \tanh \left(\frac{\beta}{N} \sum_{j \neq i} \sum_{\mu} \xi_i^{\mu} \xi_j^{\mu} S_j \right) \quad (66)$$

となる。

1.11.3 平均場近似という考え方

さて、ここで我々はこの神経回路網の性能 (最大記憶パターン数) を調べたいわけであるが、上式 (66) の右辺の \tanh の中に出てくる S_j は依然として確率変数である。従って、この変数 S_j はその平均値 $\langle S_j \rangle$ とそこからのずれ δS_j を用いて次のように表すことができる。

$$S_j = \langle S_j \rangle + \delta S_j \quad (67)$$

平均値からのずれがどの程度なのかを評価するためには具体的に (64) に従って作られる多数のアンサンブル $\{S\}$ からその統計的性質を評価することによって初めてわかるのではあるが、ここでは思い切って平均値からのずれを完全に無視してしまう近似を行おう。この近似がどの程度良いものであるか、は後に改めて議論することにする。

つまり、ここでの近似は確率変数である S_j ($j = 1, \dots, i-1, i+1, \dots, N$) の神経素子を

$$S_j = \langle S_j \rangle \quad (68)$$

と置いてしまうことである。この手の近似法を平均場近似と呼ぶ。

さて、この近似の下で (66) 式は

$$\langle S_i \rangle = \tanh \left(\frac{\beta}{N} \sum_{j \neq i} \sum_{\mu} \xi_i^{\mu} \xi_j^{\mu} \langle S_j \rangle \right) \quad (69)$$

となる。

1.11.4 局所的重なりの数値解法：平均場アニーリング

我々は各素子の状態 S_j をその平均値 $\langle S_j \rangle$ で置き換えてしまったわけであるから、これらを並べたベクトル \bar{S} :

$$\bar{S} \equiv (\langle S_1 \rangle, \langle S_2 \rangle, \dots, \langle S_N \rangle) \quad (70)$$

のように定義し, これを回路網の状態ベクトルとしよう. すると, 想起パターン ξ^1 と回路網の重なり m は

$$m = \frac{1}{N} \sum_{i=1}^N \xi_i^1 \langle S_i \rangle \quad (71)$$

のように書ける. 想起パターン ξ^1 としては何を選んでもよく, 先に見た計算機シミュレーションでは標準画像レナを用いた. しかし, ここでは簡単に $\xi^1 = (1, 1, \dots, 1)$ と選ぼう. このように選んだとしても一般性は損なわれないことに注意しておく. すると, 重なり (71) は

$$m = \frac{1}{N} \sum_{i=1}^N \langle S_i \rangle = \frac{1}{N} \sum_{i=1}^N m_i \quad (72)$$

と書き直すことができる. ただし, ここで

$$m_i \equiv \xi_i^1 \langle S_i \rangle = \langle S_i \rangle \quad (73)$$

として, 局所的重なり (local overlap) を定義した. すると, (69) 式はこの局所的重なりを用いて次のように書き直すことができる.

$$m_i = \tanh \left(\frac{\beta}{N} \sum_{j \neq i} \sum_{\mu} \xi_i^{\mu} \xi_j^{\mu} m_j \right) \quad (74)$$

従って, 想起の性能を評価する際には, 素子数 N , パターン数 p を与えて, 上式 (74) を全ての素子 i に対して数値的に解けば良い. その解が

$$\mathbf{m}^* = (m_1^*, m_2^*, m_3^*, \dots, m_N^*) \quad (75)$$

であったとするならば, 想起パターンと回路網の状態の重なり m は

$$m = \frac{1}{N} \sum_{i=1}^N m_i^* \quad (76)$$

でもって測ることができる.

具体的な (74) の解法としては, (74) を次のような連立漸化式:

$$m_i^{(n+1)} = \tanh \left(\frac{\beta}{N} \sum_{j \neq i} \sum_{\mu} \xi_i^{\mu} \xi_j^{\mu} m_j^{(n)} \right) \quad (77)$$

を全ての素子 i で収束条件:

$$|m_i^{(n+1)} - m_i^{(n)}| < \epsilon \quad (78)$$

(ϵ は例えば 10^{-5} のように小さな量) が満たされるまで反復的に解けば良い.

後にその理由を詳しく述べるが, 反復 (77) の過程で, ノイズレベル: $T \equiv \beta^{-1}$ を十分にゆっくりと $T: 2 \rightarrow 0$ のようにゼロに下げて行った場合, (77) の解 $\mathbf{m}^* = (m_1^*, m_2^*, \dots, m_N^*)$ は状態 S に関する次のエネルギー関数:

$$E(S) = -\frac{1}{N} \sum_{ij} \sum_{\mu} \xi_i^{\mu} \xi_j^{\mu} S_i S_j \quad (79)$$

の最小値を与える. つまり,

$$E(S) = -\frac{1}{N} \sum_{ij} \sum_{\mu} \xi_i^{\mu} \xi_j^{\mu} S_i S_j \geq -\frac{1}{N} \sum_{ij} \sum_{\mu} \xi_i^{\mu} \xi_j^{\mu} m_i^* m_j^* = E(\mathbf{m}^*) \quad (80)$$

が成り立つ。従って、反復式 (77) は見方を変えれば、そのエネルギー関数が (79) 式で与えられる組み合わせ最適化問題の一つの有効な解法であるとみなすことができるであろう。この手の解法を平均場アニーリング法 (meanfield annealing), あるいは確定的アニーリング法 (deterministic annealing) と呼ぶ。

実際にエネルギーの最小値が得られているか否かをパターン数が 1 つだけの自明な場合について見てみよう。このとき、エネルギー関数は

$$E(S) = -\frac{1}{N} \sum_{ij} (\xi_i^1 S_i)(\xi_j^1 S_j) \quad (81)$$

と簡略化される。さらに、想起パターン (ターゲットパターン) 全ての成分が 1 であるすればエネルギー関数は

$$E(S) = -\frac{1}{N} \sum_{ij} S_i S_j \quad (82)$$

と書ける。すると、このエネルギー関数の最小値とそれを与える S は自明であり、 $S = (1, 1, \dots, 1)$ (あるいはその真逆の $S = (-1, -1, \dots, -1)$) のとき最小値 $E = -N$ をとる。さて、パターン数が 1 つの場合、反復式 (77) は

$$m_i = \tanh \left(\frac{\beta}{N} \sum_{j \neq i} m_j \right) \quad (83)$$

となるが、 $T \rightarrow 0$ に近づける、つまり、 $\beta \rightarrow \infty$ と増加させると、上式 (83) の \tanh は sgn 関数に変わり

$$m_i = \text{sgn} \left(\frac{\beta}{N} \sum_{j \neq i} m_j \right) \quad (84)$$

この式 (84) を満たす解は $m = (1, 1, \dots, 1)$ のとき $1 = \text{sgn}(1)$ となり満たされるか、 $m = (-1, -1, \dots, -1)$ のとき $-1 = \text{sgn}(-1)$ となり満たされるか、のいずれかであり、この解は確かにエネルギー関数 (82) の最小値を与えている。

さて、上記のような平均場アニーリングを用いることにより、あるシステムサイズ N での重なりを調べることができると言うものの、これから得られる結果はサイズ $N \rightarrow \infty$ の極限での解析解 (厳密解) の近似解に過ぎない。しかし、逆に考えればこの平均場アニーリング法は神経回路の構造がどのようなものであれ、その近似解を得ることができるという利点がある。例えば、各ニューロンがある連結グラフの各ノードに配置されているとする。ここで考えてきたグラフは全てのノードがつながっている完全グラフであった。しかし、平均場アニーリング (77) は i, j に関する和を任意の「隣接するニューロンどうし」に対してとることにより、どのような構造の回路網にも適用できる。実は解析解が得られるのは回路網が完全グラフの場合に限られるので、これは大変に有効である。完全グラフの場合には 1 つのノードに入ってくる入力 $N - 1$ であり、 N が非常に大きいときには内部ポテンシャル等は中心極限定理から全てガウスに従う。しかし、例えば 2 次元正方形格子に素子が配置されている場合にはノードに入ってくる入力 N の数は 4 であり、この場合、内部ポテンシャルのガウス性などの好ましい性質がごとく崩れてくる。そうすると、もはや解析解を求めることは困難になるわけである。

この方法が有効となるであろう良い例を一つ挙げよう。人間の脳の回路網はそのつながり方の全貌を調べるには複雑過ぎるが、線虫の一種である *C. elegans* と呼ばれる虫の神経回路は 302 個のニューロンからなり、その全回路網図が完全に分かっている¹ (図 20 参照)。従って、平均場アニーリング法を用いれば、*C. elegans* の各ニューロンをこの回路図の各ノードに配置し、ニューロンどうしが Hebb 則でつながってい

¹ 例えば慶應義塾大学物理学科理論グループのライブラリ：
<http://www.phys.keio.ac.jp/staffs/kawamura-html/C-elegans-search.html> を参照されたい。

図 20: 線虫 *C. elegans*.

ると仮定したときに, *C. elegans* がメモリとして働くか否かを評価することができるであろう. 解析解という方向を断念しさえすれば, 格段にカバーできる対象が広がるというわけである. ただし, 注意しなければならないのは, 得られる結果は「近似解」であり, 得られる結果の吟味はされなければならない. ちなみに, 完全グラフで表現される全結合型の回路網に対してはこの平均場アニーリングは厳密な結果を与えることが知られている.

この平均場アニーリングに関しては「最適化問題」の回で詳しくみることにして, ここではこの程度で押さえておくに留めよう.

1.11.5 $p = \mathcal{O}(1)$ の場合の解析解とその相転移

パターン数 p が $\mathcal{O}(1)$ の場合には $N \rightarrow \infty$ の極限で解析的に重なる β 依存性を簡単に求めることができる. ここではそれを示しておこう. まず, 重なり m は

$$\begin{aligned}
 m &= \frac{1}{N} \sum_i \xi_i^1 \langle S_i \rangle \\
 &= \frac{1}{N} \sum_i \xi_i^1 \tanh \left(\frac{\beta}{N} \sum_i \sum_{\mu} \xi_i^{\mu} \xi_j^{\mu} \langle S_j \rangle \right) \\
 &= \frac{1}{N} \sum_i \xi_i^1 \tanh \left(\frac{\beta}{N} \xi_i^1 \sum_j \xi_j^1 \langle S_j \rangle + \frac{\beta}{N} \sum_j \sum_{\mu \neq 1} \xi_i^{\mu} \xi_j^{\mu} \langle S_j \rangle \right) \\
 &= \frac{1}{N} \sum_i \tanh \left(\beta m + \frac{\beta}{N} \xi_i^1 \sum_j \sum_{\mu \neq 1} \xi_i^{\mu} \xi_j^{\mu} \langle S_j \rangle \right) \tag{85}
 \end{aligned}$$

ここで, 上の (85) 式の $\tanh(\dots)$ の中身の第 2 項のオーダーを考えてみよう. すると, p が $\mathcal{O}(1)$ であれば

$$\frac{\beta}{N} \xi_i^1 \sum_j \sum_{\mu \neq 1} \xi_i^{\mu} \xi_j^{\mu} \langle S_j \rangle \sim \mathcal{O} \left(\frac{1}{N} \times \sqrt{N} \right) = \mathcal{O} \left(\frac{1}{\sqrt{N}} \right) \rightarrow 0 \tag{86}$$

となり, 神経素子数 N が無限大の極限で無視することができる. よって, 我々は重なり m に関する方程式を得ることができて

$$m = \frac{1}{N} \sum_{i=1}^N \tanh(\beta m) = \tanh(\beta m) \tag{87}$$

がその方程式である. 従って, ノイズレベル $T \equiv \beta^{-1}$ を様々に変えて (87) を解き, m と T の関係を図示すればよい. 図 21 より, ノイズ $T = 0$ の場合には完全な想起が実現されているが, ノイズの増加とともに重

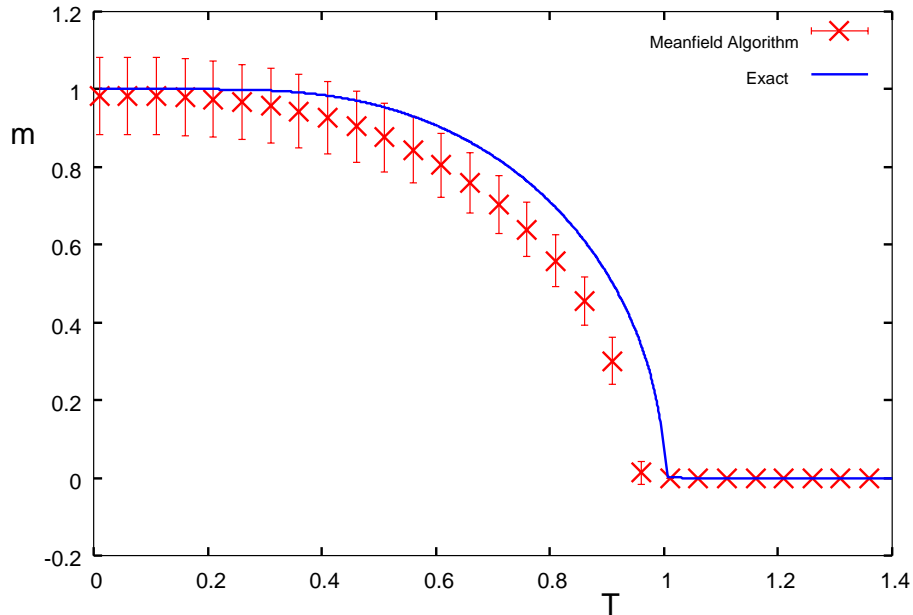


図 21: $p = O(1)$ の場合の重なり m のノイズレベル T 依存性. なお, $N = 400, p = 4$ の場合に関する平均場アルゴリズム (式 (77)). なお, ここでは各ノイズレベルで T は固定してあり, アニールはしていない) による結果も合わせてプロットしてある. 図中のエラーバーは 100 回の独立試行から算出した.

なりも減少し, $T = 1$ で重なりはゼロになる. 重なりが有限であるということは, 回路の状態ベクトル S と想起パターン ξ^1 の間には $O(N)$ 成分の一致が見られるということであり, 重なりがゼロの場合にはこの一致が $O(1)$ 成分になっているということで, $m \neq 0$ と $m = 0$ とは本質的に異なり, 前者を「想起相」後者を「非想起相」と呼ぶとすれば, ノイズが $T = 1$ になるにつれてシステムは「想起相」から「非想起相」に変化し, 回路網は連想記憶装置としては働かなくなる. この種の変化を相転移と呼ぶ. ノイズが小さく, T が小さな領域では個々の神経細胞は想起パターン方向を向いた方がエネルギー的に安定である. つまり

$$E(S) = -\frac{1}{N} \sum_{ij} \xi_i^1 \xi_j^1 S_i S_j = -\frac{1}{N} \sum_{ij} (\xi_i^1 S_i)(\xi_j^1 S_j) \tag{88}$$

であるから, 全ての素子 ($i = 1, \dots, N$) で $\xi_i^1 S_i = 1$ となればエネルギーが最小となることはすぐに見て取れる. しかし, ノイズが高くなってくると系を安定化するために個々の神経素子はバラバラな状態を取ろうとする. このバラバラな状態の程度を表す量としてエントロピー: S (回路の状態と同じ記号であるが, ここから添え字無し) の非ベクトル表示はエントロピーを表すことに注意!) を導入すれば, ノイズが大きい場合, 神経系はこのエントロピー S を大きくするように動く. そして任意のノイズレベル T では

$$F = E - TS \tag{89}$$

を最小にするように神経系は変化することになる。この F を自由エネルギーと呼ぶ。この式からノイズ $T = 0$ ではエネルギー E を最小にするように系は動き、 T が増加するにつれ、 F を最小とするためには S を大きくとるのが神経系にとっては得策になってくる²。相転移とは E を小さくしようとする、つまり、想起ベクトルの方に揃おうとする効果と S を大きくしようとするバラバラになろうとする効果が拮抗することによって生じる。

ところで、 $T = 1$ での相転移は重なり m がこの相転移点 $T = T_c = 1$ で連続的に有限値からゼロに変化する「2 次の相転移」と呼ばれる相転移であり、この場合には相転移点 T_c 及び、その相転移点近傍での重なりの振る舞いを解析的に求めることができる。まず、転移点を求めるには方程式 $m = \tanh(\beta m)$ を $m = 0$ のまわりで 1 次まで展開すると $m = \beta m$ が得られるので、これから直ちに $\beta = 1$ 、つまり、 $T = T_c = 1$ が得られる。また、 $T = T_c = 1$ 近傍での重なりの振る舞いは $m = \tanh(\beta m)$ を $m = 0$ のまわりに 3 次まで展開すれば

$$m = \beta m - \frac{1}{3}\beta^3 m^3 \tag{90}$$

となるから、これを m について解き、 $\beta - \beta_c = \beta - 1 \equiv \Delta\beta$ を用いて書き直すと

$$m^2 = \frac{3(\beta - 1)}{\beta^3} = \frac{3\Delta\beta}{(1 + \Delta\beta)^3} \simeq \Delta\beta \tag{91}$$

となり、従って、 $\Delta\beta = (1 - T)/T$ より

$$m \sim (1 - T)^{\frac{1}{2}} \tag{92}$$

のように振る舞い、 $T = T_c = 1$ で重なりはゼロに向かうことがわかる。

問 4: 平均場アニーリング (77) 式を具体的に計算機上で有限のサイズ (例えば $N = 1000$) に対して数値的に解くことにより、結果を解析解 (図 21) と比較せよ。このとき、例えば $T = 0.5$ のときの重なり m を求める際には (77) 式の反復の過程での $T = \beta^{-1}$ を $T = 2$ 程度から $T = 0.5$ に徐々に下げていくと良い。このとき、サイズを $N = 50, 100, 1000, \dots$ のように増やして行き、それぞれの結果と解析解とのズレを評価せよ。

1.11.6 $p = \mathcal{O}(N)$ の場合の解析解とその相転移

この場合には $p = \mathcal{O}(1)$ のときに $\mathcal{O}(1/\sqrt{N})$ として無視した項が $\mathcal{O}(1)$ の量となるので、この項もきちんと扱わなければならない。埋め込んだ ν 番目のパターンと神経回路網の状態との重なり m_ν は

$$\begin{aligned} m_\nu &= \frac{1}{N} \sum_i \xi_i^\nu \langle S_i \rangle \\ &= \frac{1}{N} \sum_i \xi_i^\nu \tanh \left(\frac{\beta}{N} \sum_{j \neq i} \sum_\mu \xi_i^\mu \xi_j^\mu \langle S_j \rangle \right) \\ &= \frac{1}{N} \sum_i \xi_i^\nu \tanh \left(\beta \sum_\mu \xi_i^\mu \frac{1}{N} \sum_j \xi_j^\mu \langle S_j \rangle \right) \\ &= \frac{1}{N} \sum_i \xi_i^\nu \tanh \left(\beta \sum_\mu \xi_i^\mu m_\mu \right) \end{aligned}$$

² ここまで書くと S はどのような形なのかが知りたくなる。 S を導くためにはまだ準備不足なのであるが、ここでは形だけを書いておけば $S = N \log 2 \cosh(\beta m) - (Nm/T) \tanh(\beta m)$ である。 $T \rightarrow \infty$ ($\beta = 0$) での極限では $S = N \log 2 = \log 2^N$ となる。エントロピーは「場合の数の対数」として定義されるので、この結果は「ノイズ無限大では神経素子はバラバラで全ての状態 2^N 通りを取りうるであろう」という直観と合っている。この詳細は後に見ることになる。

$$= \frac{1}{N} \sum_i \xi_i^\nu \xi_i^1 \tanh \left(\beta \xi_i^1 \sum_\mu \xi_i^\mu m_\mu \right) \quad (93)$$

となる。ここで、最後の行の変形では $\xi = \pm 1$ に対する \tanh の性質

$$\tanh(x) = \xi \tanh(\xi x) \quad (94)$$

を用いた。

次に \tanh の中を $\mu = 1$ の部分, $\mu = \nu$ の部分, それ以外の部分の 3 つの部分に分けると

$$\beta \xi_i^1 \left(\xi_i^1 m_1 + \xi_i^\nu m_\nu + \sum_{\mu \neq 1, \nu} \xi_i^\mu m_\mu \right) = \beta m_1 + \beta \xi_i^1 \xi_i^\nu m_\nu + \sum_{\mu \neq 1, \nu} \xi_i^\mu m_\mu \quad (95)$$

こうすると, (93) は

$$m_\nu = \frac{1}{N} \sum_i \xi_i^\nu \xi_i^1 \tanh \left(\beta m_1 + \beta \xi_i^1 \xi_i^\nu m_\nu + \sum_{\mu \neq 1, \nu} \xi_i^\mu m_\mu \right) \quad (96)$$

となる。

次に 1 番目のパターンを思い出している状況を考え, $m_1 \sim \mathcal{O}(1)$, $m_\nu \sim \mathcal{O}(1/\sqrt{N})$, $m_\mu \sim \mathcal{O}(1/\sqrt{N})$ ($m_\mu \neq 1, \nu$) とすると, (96) 式の右辺の \tanh の中の第 1 項は $\mathcal{O}(1)$, 第 2 項は $\mathcal{O}(1/\sqrt{N})$, 第 3 項は

$$\sum_{\mu \neq 1, \nu} \xi_i^\mu m_\mu \sim \mathcal{O} \left(\frac{1}{\sqrt{N}} \times \sqrt{p-2} \right) \sim \mathcal{O}(1) \quad (97)$$

のオーダーとなる。従って, $N \rightarrow \infty$ の素子数無限大の極限で第 2 項は第 1 項, 第 3 項に比べて無視できる。この第 2 項を微小量として展開する。

$$\frac{d}{dx} \tanh(x) = 1 - \tanh^2(x) \quad (98)$$

の関係に注意すると

$$\begin{aligned} m_\nu &= \frac{1}{N} \sum_i \xi_i^\nu \xi_i^1 \tanh \left[\beta \left(m_1 + \sum_{\mu \neq 1, \nu} \xi_i^\mu \xi_i^1 m_\mu \right) \right] \\ &+ \frac{\beta}{N} \sum_i \left\{ 1 - \tanh^2 \left[\beta \left(m_1 + \sum_{\mu \neq 1, \nu} \xi_i^\mu \xi_i^1 m_\mu \right) \right] \right\} m_\nu \\ &= \frac{1}{N} \sum_i \xi_i^\nu \xi_i^1 \tanh \left[\beta \left(m_1 + \sum_{\mu \neq 1, \nu} \xi_i^\mu \xi_i^1 m_\mu \right) \right] + \beta m_\nu \\ &- \frac{\beta}{N} \sum_i \tanh^2 \left[\beta \left(m_1 + \sum_{\mu \neq \nu, 1} \xi_i^\mu \xi_i^1 m_\mu \right) \right] m_\nu \end{aligned} \quad (99)$$

となる。ここで $\sum_{\mu \neq \nu, 1} \xi_i^\mu \xi_i^1 m_\mu$ は $\xi_i^\mu \xi_i^1$ が ± 1 をとることから, 平均ゼロ, 分散

$$\sum_{\mu \neq 1} (\xi_i^\mu \xi_i^1)^2 m_\mu^2 = \sum_{\mu \neq \nu, 1} m_\mu^2 \equiv \alpha r \quad (100)$$

のガウス分布に従う³. ここで $N \rightarrow \infty$ では次の性質:

$$\frac{1}{N} \sum_i f(x_i) = \int_{-\infty}^{\infty} dx p(x) f(x) \quad (101)$$

が成り立つことに注意する. この性質を統計量 $f(x)$ の自己平均性 (self-averaging) と言う. ここで $p(x)$ は確率変数 x の従う分布である. つまり, 確率変数 x_i によって本来は $f(x_i)$ 自体もばらつくはずなのであるが, 十分に大きなシステムを考えると $f(x_i)$ はその平均値: $\int_{-\infty}^{\infty} dx p(x) f(x)$ に一致する. $N \rightarrow \infty$ の極限では $f(x_i)$ 自分自身がその平均値に一致するわけである (だから「自己平均」と呼ぶ).

これは直観的には次のようにして理解できる. 今の場合 $x_i = \sum_{\mu \neq \nu, 1} \xi_i^\mu \xi_i^1 m_\mu$ は ξ^μ を様々変えれば変わるのだから, その関数である $f(x_i)$ も変わる. もちろん, レナの 2 値化画像を ξ^μ に選んだ場合と, 印鑑のような 2 値画像を ξ^μ に選んだ場合とでは $f(x_i)$ は異なる. 従って, 我々は全ての可能な ξ^μ に関して $f(x_i)$ を平均したものを求めなければならない. しかし, 十分に大きな N を考えた場合, つまり, 十分に大きなサイズの画像を考え, それを小さな部分画像に分割してみると, その十分大きな画像の中にはレナの 2 値化画像, 印鑑の画像はおろか, ありとあらゆる全ての 2 値画像がその部分画像としてその「どこかには」含まれることになると考えてよいであろう (これはちょっと奇異に聞こえるかもしれない. 大きさ無限大の画像などというものは我々の想像力を超えている. しかし, これは事実である). 従って, 十分大きな画像 ξ^μ を考え, その $f(x_i)$ を測定すれば, それは実質全ての可能な部分画像に対して平均操作を行ったものに等しくなると予想できる.

さて, この自己平均性を用いると

$$\begin{aligned} \frac{1}{N} \sum_i \tanh^2 \left[\beta \left(m_1 + \sum_{\mu \neq \nu, 1} \xi_i^\mu \xi_i^1 m_\mu \right) \right] &= \int_{-\infty}^{\infty} \frac{dx}{\sqrt{2\pi\alpha r}} e^{-\frac{x^2}{2\alpha r}} \tanh^2[\beta(m_1 + x)] \\ &= \int_{-\infty}^{\infty} Dz \tanh^2[\beta(m_1 + \sqrt{\alpha r}z)] \equiv q \end{aligned} \quad (102)$$

となる. ここで $Dz \equiv dz e^{-\frac{z^2}{2}} / \sqrt{2\pi}$ でガウス積分測度を定義した.

以上を (99) 式に代入すると,

$$m_\nu = \frac{1}{N} \sum_i \xi_i^\nu \xi_i^1 \tanh \left[\beta \left(m_1 + \sum_{\mu \neq 1, \nu} \xi_i^\mu \xi_i^1 m_\mu \right) \right] + \beta m_\nu - \beta q m_\nu \quad (103)$$

となるので, これを m_ν に関して解いて

$$m_\nu = \frac{1}{[1 - \beta(1 - q)]} \frac{1}{N} \sum_i \xi_i^\nu \xi_i^1 \tanh \left[\beta \left(m_1 + \sum_{\mu \neq 1, \nu} \xi_i^\mu \xi_i^1 m_\mu \right) \right] \quad (104)$$

を得る.

さて, q の表式の中には r や m_1 が現れるので, これらを自己無撞着 (self-consistent) に求めなくてはならない. そこで, まずは上式の両辺の自乗を計算すると

$$\begin{aligned} m_\nu^2 &= \left[\frac{1}{1 - \beta(1 - q)} \right]^2 \frac{1}{N^2} \sum_{ij} \xi_i^\nu \xi_i^1 \xi_j^\nu \xi_j^1 \\ &\quad \times \tanh \left[\beta \left(m_1 + \sum_{\mu \neq 1, \nu} \xi_i^\mu \xi_i^1 m_\mu \right) \right] \tanh \left[\beta \left(m_1 + \sum_{\mu \neq 1, \nu} \xi_j^\mu \xi_j^1 m_\mu \right) \right] \end{aligned} \quad (105)$$

³ ここで $m_\mu \sim \mathcal{O}(1/\sqrt{N})$ より, $m_\mu^2 \sim \mathcal{O}(1/N)$ となり

$$\sum_{\mu \neq 1, \nu} m_\mu^2 \sim \mathcal{O}((p-2)/N) \sim \mathcal{O}(\alpha)$$

であるから r はオーダー 1 の量になることに注意.

となるが, 生き残るのは $i = j$ を満たす項のみだから

$$m_\nu^2 = \left[\frac{1}{1 - \beta(1 - q)} \right]^2 \frac{1}{N^2} \sum_i \tanh^2 \left[\beta \left(m_1 + \sum_{\mu \neq 1, \nu} \xi_i^\mu \xi_i^1 m_\mu \right) \right] \quad (106)$$

が得られる. 従って r は

$$\begin{aligned} r = \frac{1}{\alpha} \sum_{\nu \neq 1} m_\nu^2 &= \frac{1}{\alpha} \times (N\alpha - 1) \times \frac{1}{N^2} \left[\frac{1}{1 - \beta(1 - q)} \right]^2 \sum_i \tanh^2 \left[\beta \left(m_1 + \sum_{\mu \neq 1, \nu} \xi_i^\mu \xi_i^1 m_\mu \right) \right] \\ &= \left[\frac{1}{1 - \beta(1 - q)} \right]^2 \frac{1}{N} \sum_i \tanh^2 \left[\beta \left(m_1 + \sum_{\mu \neq 1, \nu} \xi_i^\mu \xi_i^1 m_\mu \right) \right] \\ &= \left[\frac{1}{1 - \beta(1 - q)} \right]^2 \int_{-\infty}^{\infty} Dz \tanh^2[\beta(m_1 + \sqrt{\alpha r}z)] \\ &= \frac{q}{[1 - \beta(1 - q)]^2} \end{aligned} \quad (107)$$

を得る.

これで q, r が求まった. 残るは m_1 であるが, これは

$$m_\nu = \frac{1}{N} \sum_i \xi_i \tanh(\beta \sum_\mu \xi_i^\mu m_\mu) \quad (108)$$

で $\nu = 1$ とおいて,

$$\begin{aligned} m_1 &= \frac{1}{N} \sum_i \xi_i^1 \tanh(\beta \sum_\mu \xi_i^\mu m_\mu) \\ &= \frac{1}{N} \sum_i (\xi_i^1)^2 \tanh(\beta \sum_\mu \xi_i^1 \xi_i^\mu m_\mu) \\ &= \frac{1}{N} \sum_i \tanh \left(\beta m_1 + \beta \sum_{\mu \neq \nu=1} \xi_i^1 \xi_i^\mu m_\mu \right) \\ &= \int_{-\infty}^{\infty} Dz \tanh[\beta(m_1 + \sqrt{\alpha r}z)] \end{aligned} \quad (109)$$

が得られる. 以上をまとめると ($m_1 = m$ とする)

$$m = \int_{-\infty}^{\infty} Dz \tanh[\beta(m + \sqrt{\alpha r}z)] \quad (110)$$

$$q = \int_{-\infty}^{\infty} Dz \tanh^2[\beta(m + \sqrt{\alpha r}z)] \quad (111)$$

$$r = \frac{q}{[1 - \beta(1 - q)]^2} \quad (112)$$

となる.

従って, あるノイズレベル β のときの最大記憶パターンを求めるには, 上の 3 本の連立方程式を解き, $m \neq 0$ を解に持つような $\alpha = \alpha_{\max}$ を求めれば

$$p_{\max} = \alpha_{\max} N \quad (113)$$

として決定される.

(A) ノイズゼロの場合の記憶容量

まずはノイズゼロ: $T = 0$ の場合 $\beta \rightarrow \infty$ に関して最大記憶パターンを求めてみることにしよう.

$\beta \rightarrow \infty$ の極限で m は

$$m = \int_{-\infty}^{\infty} Dz \operatorname{sgn}(m + \sqrt{\alpha r} z) = \int_{-\frac{m}{\sqrt{\alpha r}}}^{\infty} Dz - \int_{-\infty}^{-\frac{m}{\sqrt{\alpha r}}} Dz = 1 - 2H\left(\frac{m}{\sqrt{\alpha r}}\right) \quad (114)$$

となる. ここで関数 $H(x)$ を

$$H(x) \equiv \int_x^{\infty} Dz \quad (115)$$

で定義した.

$(1 - q)$ は

$$\begin{aligned} 1 - q &= \int_{-\infty}^{\infty} Dz (1 - \tanh^2[\beta(m + \sqrt{\alpha r} z)]) \\ &\simeq \frac{1}{\sqrt{2\pi}} e^{-\frac{m^2}{2\alpha r}} \int_{-\infty}^{\infty} dz (1 - \tanh^2[\beta(m + \sqrt{\alpha r} z)]) \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{m^2}{2\alpha r}} \int_{-\infty}^{\infty} \frac{d}{dz} \left(\frac{\tanh[\beta(m + \sqrt{\alpha r} z)]}{\beta\sqrt{\alpha r}} \right) dz \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{m^2}{2\alpha r}} \left[\frac{\tanh[\beta(m + \sqrt{\alpha r} z)]}{\beta\sqrt{\alpha r}} \right]_{-\infty}^{\infty} \\ &= \frac{1}{\beta} \sqrt{\frac{2}{\pi\alpha r}} e^{-\frac{m^2}{2\alpha r}} \end{aligned} \quad (116)$$

となる. 従って, $C \equiv \beta(1 - q)$ とおくと, 上 3 式は

$$C = \sqrt{\frac{2}{\pi\alpha r}} e^{-\frac{m^2}{2\alpha r}} \quad (117)$$

$$r = \frac{1}{(1 - C)^2} \quad (118)$$

$$m = 1 - 2H\left(\frac{m}{\sqrt{2\alpha r}}\right) \quad (119)$$

となり, さらに

$$y \equiv \frac{m}{\sqrt{2\alpha r}} \quad (120)$$

で新しい変数 y を導入すると, 上式はさらに簡略化され

$$y \left(\sqrt{2\alpha} + \frac{2}{\sqrt{\pi}} e^{-y^2} \right) = 1 - 2H(y) \quad (121)$$

が得られる. これは, 図 22 のようにグラフを用いて解くことができる. 最終的に, 最大記憶パターンは $\alpha_{\max} \sim 0.138$ となるので,

最大記憶パターン数 : $p_{\max} = 0.138N$

が求まる.

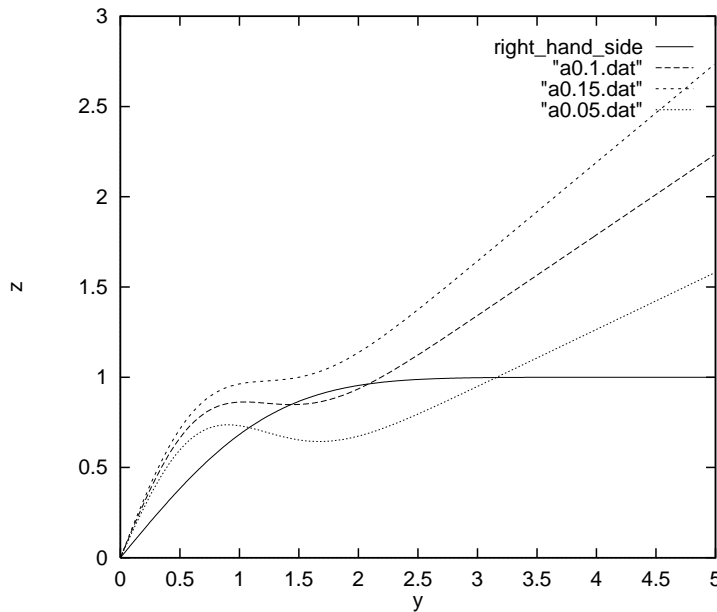


図 22: 実線が (121) の右辺, つまり $1 - 2H(y)$ である. 破線, 点線が (121) の左辺であり, これは α に依存して変化する. α を変化させて行くと, およそ $\alpha = 0.138$ で両者の交点がなくなる.

(B) 有限ノイズの場合の記憶容量と相関

話が長くなってしまったので, 簡単にではあるがこの辺りでこれまでの問題を整理しておくことにしよう. 我々の問題は, 神経素子の出力に確率的な動作を導入した場合, つまり

$$P(S_i = 1) = \frac{e^{\beta h_i}}{e^{\beta h_i} + e^{-\beta h_i}}$$

$$P(S_i = -1) = \frac{e^{-\beta h_i}}{e^{\beta h_i} + e^{-\beta h_i}} = 1 - P(S_i = 1)$$

$$h_i \equiv \frac{1}{N} \sum_{j \neq i} \sum_{\mu=1}^p \xi_i^\mu \xi_j^\mu S_j$$

で与えられる確率により, 全ての神経素子 $S_i (i = 1, \dots, N)$ が非同期的状態変更をした場合, 任意の記憶パターン $\xi^1 = (\xi_1^1, \xi_2^1, \dots, \xi_N^1)$ と神経回路の状態 $S = (S_1, S_2, \dots, S_N)$ との間の重なり;

$$m = \frac{1}{N} \sum_{i=1}^N \xi_i^1 S_i$$

が有限となるような最大の記憶パターン数 $p_{\max} = \alpha N$ を求めることであった. 最大記憶パターン数⁴ α は回路網の確率的な動作の程度を表すノイズレベル: $T = \beta^{-1}$ に依存することに注意しよう. 明らかに $T = 0$ の場合にはノイズフリーな回路網になり, T を大きくして行けば個々の神経素子の挙動は不確かさを増す ($T \rightarrow \infty$ で完全にランダムに ± 1 を出力する素子になる). よって, 有限ノイズの連想記憶装置の性能を評価するためには記憶させたパターン数とノイズレベル, すなわち (α, T) の組をある値に固定したときに, 重なり m が有限の値を持つかどうかを調べればよい. もし, $m \neq 0$ の解を持つならば回路網はこのノイズレベル T の下で αN 個のパターンを記憶する連想記憶装置として動作することが結論付けられる.

⁴ 正確に言えばパターン数は α に全素子数 N をかけた αN であるが, 実質的には α を決めればパターン数は決まるわけであるから, 以後, α をもって「パターン数」と呼ぶ場合があることに注意されたい.

さて, (α, T) の一組が与えられたときに, 重なり m を含む回路網は次の状態方程式で記述された.

$$m = \int_{-\infty}^{\infty} Dz \tanh[\beta(m + \sqrt{\alpha r}z)] \quad (122)$$

$$q = \int_{-\infty}^{\infty} Dz \tanh^2[\beta(m + \sqrt{\alpha r}z)] \quad (123)$$

$$r = \frac{q}{[1 - \beta(1 - q)]^2} \quad (124)$$

但し, ここでは積分測度を $Dz \equiv dze^{-z^2/2}/\sqrt{2\pi}$ と置いている. 上記の方程式はこれ以上簡単にはならない (z に関する積分が実行できない). 従って, 上記の方程式の中に含まれる z に関する数値積分を実行し, この連立方程式を数値的に解くことになる⁵.

さて, 式 (124) で r は q のみで書けるのであるから, (124) 式を (122)(123) 式に代入すれば, 形式的に次のような m と q の 2 元非線形連立方程式を得る.

$$m = M(m, q) \quad (125)$$

$$q = Q(m, q) \quad (126)$$

この連立方程式の解き方は様々あろうが, 例えば, 上の方程式を一種の「写像」, つまり

$$m_{t+1} = M(m_t, q_t) \quad (127)$$

$$q_{t+1} = Q(m_t, q_t) \quad (128)$$

とみなし, 適当な初期条件 (m_0, q_0) からスタートし⁶, 隣り合う時刻 $t = t, t + 1$ のときの m_t 及び q_t の差が, ある範囲内 ϵ に収まったとき, つまり

$$|m_{t+1} - m_t| < \epsilon \quad (129)$$

$$|q_{t+1} - q_t| < \epsilon \quad (130)$$

が同時に成立したときの $(m_t, q_t) = (m, q)$ を解とするのも有効な方法の一つであろう. ϵ は例えば $\epsilon = 10^{-5}$ と選ぶなど, 必要とする精度に応じて決めてやればよい. もちろん, これらは「非線形」な写像であるから, 一般的に言ってパラメータ (α, T) の取り方, 初期条件の選び方によっては振動解やカオス解が生じたりする可能性もある. しかし, 今の場合にはそのような解は現れず, 素直に固定点に収束する (その意味で非線形性はさほど強くない)⁷.

以上の方法でノイズレベルを $T = 0.4$ で固定した場合の重なり m の記憶パターン数 α 依存性を調べたので, それを図 23 に載せよう. この図より, パターン数 α を徐々に増やしていくと, ある臨界値 $\alpha_c \simeq 0.077$ で突然 m は有限の値からゼロへと変化し, これ以上のパターンを詰め込むと, 今まで思い出すことができたパターンさえも記憶から消失してしまうことがわかる. いわば「頭の中が真っ白な状態になってしまう」というわけである⁸.

⁵ 最終的には数値積分やコンピュータを用いて連立方程式を解くことになったとはいえ, これらの状態方程式を導出した時点で問題自体はすっかり解けていることに注意されたい. これは状態方程式の中に回路網のサイズ (神経素子の個数) N が入っていないことから明らかである ($N \rightarrow \infty$ の極限操作で解析解を得たのであった.) 先に述べた平均場アニーリング法で計算機を用いて数値的に重なりを評価する場合には予め素子数 N を与えなければならず, 当然, 得られる結果は $N \rightarrow \infty$ の場合の近似解である.

⁶ 写像の状態変更毎に数値積分を実行する必要があることに注意されたい.

⁷ もちろん, $M(m, q) - M, Q(m, q) - q$ の 1 回微分を用いるニュートン法で解を求めてもよい.

⁸ 先にも述べたように, 外部からの制御パラメータ (α, T) の変化によって, m などの巨視的な量 (秩序変数) に特異性が生じることを相転移と呼ぶ. また, この例のように, (「自由エネルギー」と呼ばれる物理量の制御変数に関する一回微分で得られる) m に不連続な変化があらわれる相転移を, 先に見たノイズレベルを増加していったときに m が連続的にゼロになった 2 次転移に対して 1 次の相転移 (1 次転移) と呼んでいる.

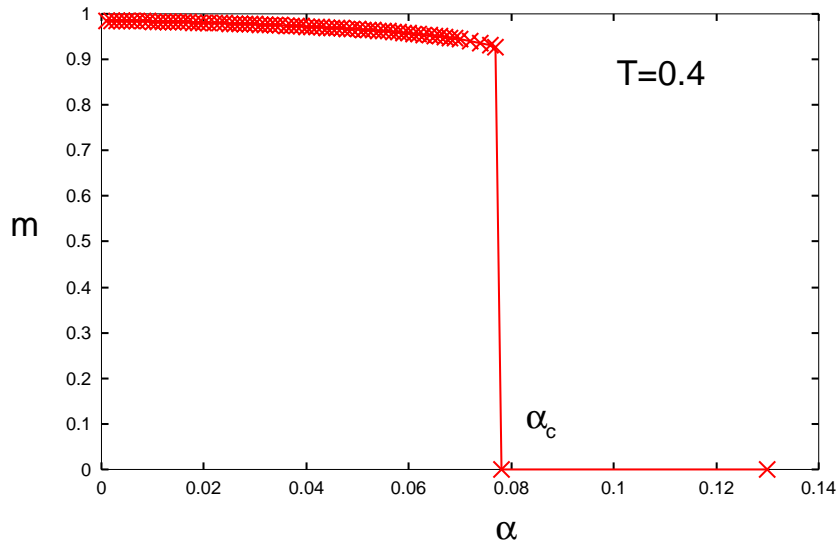


図 23: ノイズレベルを $T = 0.4$ とおいた場合の重なり m の記憶パターン数 α 依存性. $\alpha \simeq 0.077$ (図の critical point α_c) で有限の m の値から $m = 0$ へと突然変化する.

さて、この例では $T = 0.4$ の場合であったが、 T をも変えて行き、 $m \neq 0$ である領域と $m = 0$ である領域の境界を α - T 平面の中にも書き込むと、この回路網の動作がつかみ易くなる。このような図を相図 (phase diagram) と呼ぶ。これを図 24 に載せる。この図より、想起相 (retrieval) と書かれている領域の (α, T) を選べば、回路網は連想記憶装置として動作する。言い方を変え、この retrieval 相は $m \neq 0$ で特徴付けられた相である。一方、スピングラス相 (spin glass) 相では $m = 0$ であり⁹、連想記憶装置としては動作しない。retrieval 相から spin glass 相への相転移は一次転移である。つまり、retrieval 相 \rightarrow spin glass 相の過程で m が不連続に有限値からゼロに変化する (図 23 を参照されたい)。

さて、ノイズゼロ、つまり、 $T = 0$ の状況では方程式は簡略化され、比較的簡単に $\alpha_c \simeq 0.138$ という結果を得たのであった。今の場合、図 24 の $T = 0$ の軸 (横軸) と相境界の接点を見ると、確かに $T = 0$ で $\alpha_c \simeq 0.138$ となっており、我々が以前行った計算結果とつじつまが合っていることが確認できる。

さて、スピングラス相と常磁性相の相境界線は 2 次転移であるから、解析的に求めることができる。(123)(124) 式をそれぞれ、 $q = 1, m = r = 0$ のまわりで展開すれば

$$q \simeq \beta^2 \alpha r \int_{-\infty}^{\infty} D z z^2 = \beta^2 \alpha r \tag{131}$$

$$r \simeq \frac{q}{(1 - \beta)^2} \tag{132}$$

であり、この両式から r を消去して $T = \beta^{-1}$ について解けば直ちに

$$T = 1 + \sqrt{\alpha} \tag{133}$$

が得られる。これがスピングラス相と常磁性相の相境界を表す式である。

⁹ 当然の疑問として「スピングラスとは何か?」というのがあるに違いない。ここでは「スピングラスとは磁性体の一種である」と言っておくにとどめるが、回路網の状態という観点からは「スピングラス状態とは記憶してもいないパターンに回路網の状態が固定されてしまった状況」であると考えられる。詳細はこの講義で後ほど説明する。

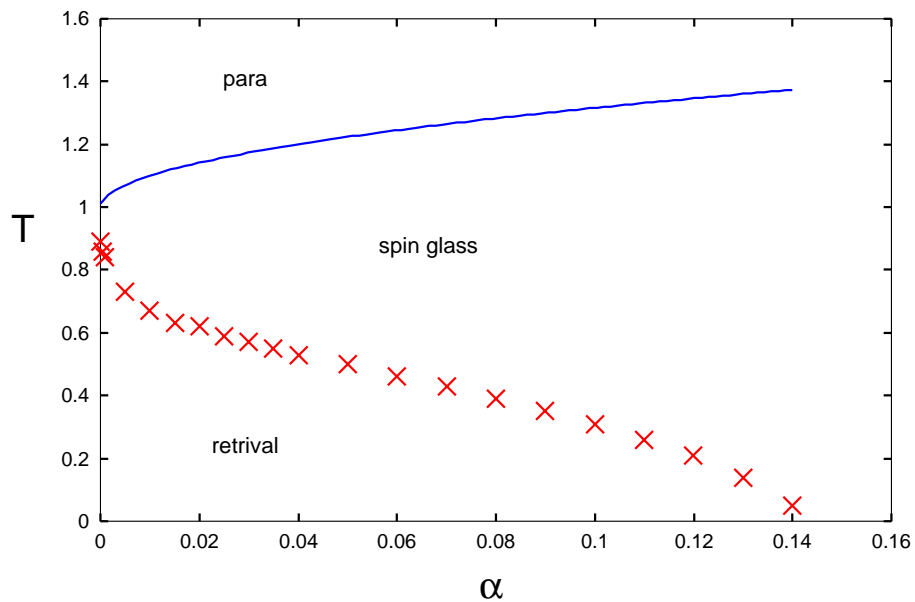


図 24: 確率的神経回路網の相図. 想起相 (retrival) と書かれている領域の (α, T) を選べば, 回路網は連想記憶装置として動作する. 一方, スピングラス相 (spin glass) では $m = 0$ であり, 連想記憶装置としては動作しない. retrival 相から spin glass 相への相転移は一次転移である. つまり, retrival 相 \rightarrow spin glass 相の過程で m が不連続に有限値からゼロに変化する (図 23 参照).