

学位請求論文

ニューラルネットワークのダイナミクス
に関する研究

指導教官 岡部 洋一 教授

東京大学大学院 工学系研究科 先端学際工学専攻

57247 掛谷 英紀

平成9年12月19日提出

目次

第1章	序論	1
1.1	ニューラルネットワーク研究の背景	1
1.2	本研究の背景と目的	2
1.3	本論文の構成	3
第2章	ニューロダイナミクスの幾何	4
2.1	形式ニューロンモデル	4
2.2	ニューロダイナミクスの幾何	5
2.3	まとめ	9
第3章	対称ネットワークの幾何	10
3.1	自己相関連想記憶の幾何	10
3.2	記憶空間の分割	35
3.3	まとめ	45
第4章	非対称ネットワークの幾何	46
4.1	非対称行列の幾何構造	46
4.2	相互相関連想記憶	48
4.3	複雑な動的振舞の解析	55
4.4	まとめ	62
第5章	ニューロウインドウ法	63
5.1	ニューロウインドウ法	63
5.2	自己相関連想記憶のニューロウインドウ法	67
5.3	相互相関連想記憶のニューロウインドウ法	98
5.4	(0,1)出力系のニューロウインドウ連想記憶	104
5.5	オートマトンにおける冗長度低減モデル	109
5.6	まとめ	124
第6章	連想記憶の統計理論解析	125
6.1	ニューロウインドウ連想記憶の有限系解析	125
6.2	ニューロウインドウ連想記憶のSCSNA	127
6.3	部分反転法の統計神経力学	134
6.4	ニューロウインドウの統計神経力学	140
6.5	まとめ	146

目次

第7章	連続時間ダイナミクス	147
7.1	連続時間ニューロダイナミクス	147
7.2	連続時間自己連想記憶	148
7.3	連続時間相互連想記憶	163
7.4	ランダムネットワークの連続時間ダイナミクス	171
7.5	まとめ	174
第8章	組み合わせ最適化の幾何とその応用	175
8.1	ニューロによる組み合わせ最適化	175
8.2	分割問題の幾何とその応用	177
8.3	巡回セールスマン問題の幾何と応用	184
8.4	まとめ	188
第9章	学習に関する考察	189
9.1	直交学習に関する考察	189
9.2	BP学習に関する考察	198
9.3	まとめ	205
第10章	結論	206
10.1	本論文のまとめ	206
10.2	今後の課題	207
	謝辞	209
	参考文献	211
第A章	離散時間対称ネットワークのエネルギー	214
第B章	ニューロウィンドウ法のSCSNA	216
第C章	部分反転法の統計神経力学	221
第D章	ニューロウィンドウ法の統計神経力学	227
	発表文献	233

第 1 章： 序論

本研究の目的は、ニューラルネットワーク（神経回路網）が有する力学的性質の数理解析、及び、数理解析で得られた理論に立脚した情報処理システムの提案である。この章では、本研究の背景をニューラルネットワーク一般論と本研究の具体論の二つの視点から論じることで本研究の意義の明確化を図る。

1.1 ニューラルネットワーク研究の背景

ニューラルネットワークの研究は、脳を模倣して新たな計算機を実現することを目標としている。脳の模倣を試みる背景には、人間にできて計算機や機械にできない情報処理が数多く残されているという現状がある。勿論、原理的には、現在のフォンノイマン型計算機でも計算可能性を満たす問題は全て解くことが可能である。しかし、その計算機アーキテクチャは論理演算をベースにしているので、所望の問題を解くためには、それを論理式に書き下すプログラミングが必要になる。であるから、問題によっては、目的の計算を実行するためのプログラムを生成することが困難となる。例えば、知能情報処理の分野における翻訳、音声認識、文字認識など、文脈依存性のあるパターン情報処理については、人間と同程度の精度を計算機で達成することは現時点では不可能である。また、運動タスクにおいても、人間と同様の運動を自動機械で実現することは困難を極める。このように、従来の計算機アーキテクチャでは実現が難しいタスクを、脳の情報処理形態に習って達成することが、ニューラルネットワーク研究の動機となっている。

しかしながら、脳を模倣するといっても、生物の脳は非常に複雑であり、それを写實的に模倣するのは非現実的である。そこで、実際には脳の計算様式のある特徴を抽出し、それを利用した計算機アーキテクチャ／アルゴリズムを試みることになる。脳神経回路の最も顕著な特徴の一つとして、それがニューロン（神経細胞）という閾値素子的な性質を持つ細胞のネットワークで構成されている点が挙げられる。そこで、この特徴を生かした計算システムの構築が試みられている。現在、コンピュータサイエンスの分野で「ニューラルネットワーク」というときには、このニューロンを真似た閾値素子（マッカロー＝ピッツの形式ニューロン [25]）を組み合わせて構成したネットワークを指すことが一般的である。本研究でも、その定義に基づくニューラルネットワークを前提に議論を行うことにする。

マッカロー＝ピッツの形式ニューロンに基づくニューラルネットワークが実際の工学的問題に応用された例は、バックプロパゲーション学習 [34] を用いた関数近似による制御／予測問題への応用、ホップフィールドネットワーク [17] 等を用いた組み合わせ最適化問題への応用の二つが主である。どちらも、どうしてもニューロでなければできないという類いのタスクではなく、上記で述べた人間は得意で計算機が不得意なタスクを実

現しているとは言難い面がある。このように、ニューラルネットワークの応用が進んでいない要因としては、その系の複雑性が挙げられる。ニューラルネットワークは、非線形多次元システムであるので、一般論としてその理論解析を進めることは容易ではない。そのため、ニューラルネットワーク研究の現状は、理論的な取り扱いを容易にするために理想化された系に関する数理解析と、種々の具体的な問題に特化したシステムのチューニングに二極化する傾向にある。その結果、数理的な議論の積み上げからは実際の問題に応用できるようなアルゴリズムの提示が難しく、一方応用ではニューロをブラックボックスとして使うしかないという状況が生じている。

このように、ある意味で閉塞状態ともいえるニューロ研究の状況において、研究の新たな展開としては次のような方向性が考えられる。まず、第一に、閾値素子のネットワークとしてのニューロは、一般的にどのような性質を持つかを洗い直し、実用上ニューラルネットワークを使用するメリットは何なのかを明確化する必要がある。そのためには、今までのように単に理想化された単純な系の理論解析ではなく、ニューロで実際に情報処理システムの作成を考える上で有用な理論構築が要求される。次に、これまでのニューロの枠組みから踏み出すため、生物の情報処理系から「閾値素子のネットワーク」という形式以上の特徴的情報処理様式の抽出を行い、それを工学システムとして実現できるように定式化していく必要がある。

本論文では、この二つの課題のうち、前者について取り上げることにする。後者の課題はより大きな課題であるが、本論文では対象外とする。後者の視点からの研究が強く要求される時期にきているという現状認識を筆者は持っているものの、著者の力不足によりその課題を追求するに至らなかった次第である。

1.2 本研究の背景と目的

上で述べたとおり、本研究は閾値素子のネットワークとして定式化される現在のニューラルネットワークの枠組で計算システムを組むことの意義を明確化することを目標とする。

本研究の背景として、金道による連想記憶の研究 [19, 20] がある。金道はニューロダイナミクスを超球面上のフローとして幾何的に捕らえることで、自己相関連想記憶のダイナミクスを説明した。著者は修士論文において、この解析法を応用し、ニューロウインドウ連想記憶を提案した。この幾何的解析手法は、連想記憶に限らずニューロダイナミクス全般に応用可能であると考えられる。但し、金道の提案した解析法は荷重行列が対称行列であることを前提にしていたので、非対称行列を含む一般論を行うためには解析手法の拡張が必要である。また、より広い応用モデルを扱うには、離散時間ダイナミクスだけでなく連続時間ダイナミクスにも理論を拡張する必要がある。更に、幾何的手法は定量的な解析を行うには不十分であるので、定量的な議論を行うために統計手法を用いた解析を並列に行っていくことも必要となる。

本論文においては、以上の課題を踏まえ、ニューロダイナミクスの一般に関する理論構築のため、ニューロダイナミクスの幾何理論を非対称結合系や連続時間系に拡張する

ことを目指す。また、幾何理論の定量性の不足を補うため、幾何理論をもとに得られたシステムについての統計理論を構築し、その定量的解析を試みる。更に、ニューロダイナミクス一般について得られた理論的指針をもとに、自己相関連想記憶に限らず、相互相関連想記憶、組み合わせ最適化、ランダムネットワーク、リカレントネットワーク等、ニューラルネットワーク一般を対象に新たなモデルやアルゴリズム等の提案を行うことを目指す。

1.3 本論文の構成

本論文の構成は以下の通りである。まず、第2章において本論文の議論の基礎となるニューロダイナミクスの幾何的描像について述べる。この幾何理論をもとに、第3章においては対称結合を持つニューラルネットワークの性質を、第4章においては非対称結合を持つニューラルネットワークの性質を議論する。第5章においては、ニューロウィンドウ法を用いた連想記憶における選択的想起モデルを提案する。第6章においては、提案したニューロウィンドウ連想記憶について統計理論を用いて定量的解析を行う。第7章では、連続時間ニューロダイナミクスを離散時間モデルと比較しながら議論する。第8章においては、組み合わせ最適化問題を解くニューラルネットワークを幾何的な視点から解析し、並列デジタル計算機で高速に準最適解を得るアルゴリズムを提案する。最後に、第9章で、直交学習やBP学習などの繰り返し学習について簡単な考察を加える。

第 2 章： ニューロダイナミクスの幾何

ニューラルネットワークは非線形多次元システムであり、そのダイナミクスの挙動は複雑である。本章においては、ニューラルネットワークのダイナミクスの解析に幾何的視点を導入し、ニューロダイナミクスの特徴を明確化する。本章の内容はニューロダイナミクスを扱う本論文における議論のベースとなるものである。

まず、第1節では、現在の Artificial Neural Network (人工神経回路網) の構成要素になっている形式ニューロンモデルについて説明する。そして、第2節では、形式ニューロンモデルより構成されるニューラルネットワークの力学系の幾何構造を与える。

2.1 形式ニューロンモデル

現在、ニューラルネットワーク (神経回路網) について多種多様な研究が行われているが、これらの研究は McCulloch と Pitts によるニューロン (神経細胞) モデルがその基本になっている [25]。McCulloch と Pitts の形式ニューロンモデルとは、ニューロンの膜電位 u が入力 x_i の重みつき和

$$u = \sum_{i=1}^N w_i x_i \quad (2.1)$$

で定まり、膜電位 u が閾値 h を越えるとそのニューロンは 1 を出力し (発火状態)、越えなければ 0 を出力する (非発火状態) とするモデルである。このモデルの入出力関係を式で表すと

$$z = 1 \left(\sum_{i=1}^N w_i x_i - h \right) \quad (2.2)$$

となる。ここで、 z はニューロンの出力値であり、 $1(u)$ はステップ関数

$$1(u) = \begin{cases} 1, & \text{if } u \geq 0 \\ 0, & \text{if } u < 0 \end{cases} \quad (2.3)$$

である。

その後、現在までの研究において、このモデルを基本にして、いくつかのニューロンモデルのバリエーションが考えられている。例えば、数理的に扱いやすいモデルとしては、ステップ関数の代わりに符号関数

$$\text{sgn}(u) = \begin{cases} 1, & \text{if } u \geq 0 \\ -1, & \text{if } u < 0 \end{cases} \quad (2.4)$$

を用いるものがある。また、ニューロンの出力が連続値を取るモデルとしては、ヘビサイド関数

$$f(u_i) = \frac{1}{1 + \exp(-\beta u_i)} \quad (2.5)$$

やシグモイド関数

$$g(u_i) = \frac{1 - \exp(-\beta u_i)}{1 + \exp(-\beta u_i)} \quad (2.6)$$

などが一般的に使用されている。ここで、 β は関数の形を決定する定数で、 $\beta = \infty$ の時、 $f(u) = 1(u)$, $g(u) = \text{sgn}(u)$ になる。更に、時間軸を考慮に入れると、膜電位 u が

$$\tau \frac{du_i}{dt} = -u_i + \sum_{j=1}^N w_{ij} x_j \quad (2.7)$$

のように連続的に変化するモデルも考えることができる。本論文においては、以上に挙げたニューロンモデルを場合によって使い分けて議論を進めることにする。

2.2 ニューロダイナミクスの幾何

前節で述べたように、ニューロンモデルには様々なものがあり、どのニューロンモデルを用いるかによって、ニューラルネットワークのダイナミクスの挙動も変わってくる。この節では、中でもそのダイナミクスが最も簡単な幾何構造をもつ符号関数ニューロンを使った離散時間同期型ネットワークを最初にとりあげる。その他のニューロンモデルを用いたニューラルネットワークについては、符号関数を用いた離散時間同期モデルでの議論をベースにし、それを拡張する形でダイナミクスの幾何構造を論じることにする。

2.2.1 離散時間同期ダイナミクスの幾何

ニューロンの活性化関数として符号関数を用いる離散時間同期型ネットワークのダイナミクスは以下のように定式化される。まず、神経回路網は N 個のニューロンよりなるとする。各時刻 t における i 番目のニューロンの状態を $x_i(t)$ と記し、ネットワークの状態をベクトル $\mathbf{x}(t) = (x_1(t), x_2(t), \dots, x_N(t))^T$ で表現する。この時、離散時間同期ダイナミクスは

$$x_i(t+1) = \text{sgn}\left(\sum_{j=1}^N w_{ij} x_j(t)\right) \quad (2.8)$$

$$\text{sgn}(u) = \begin{cases} 1, & \text{if } u \geq 0 \\ -1, & \text{if } u < 0 \end{cases} \quad (2.9)$$

で表される時間発展方程式をもつ。

式(2.8)で表されるニューラルネットワークのダイナミクスは次のように二段階に分けて考えることができる。まず、第一段階として、状態ベクトル $\mathbf{x}(t)$ は、記憶行列 W に

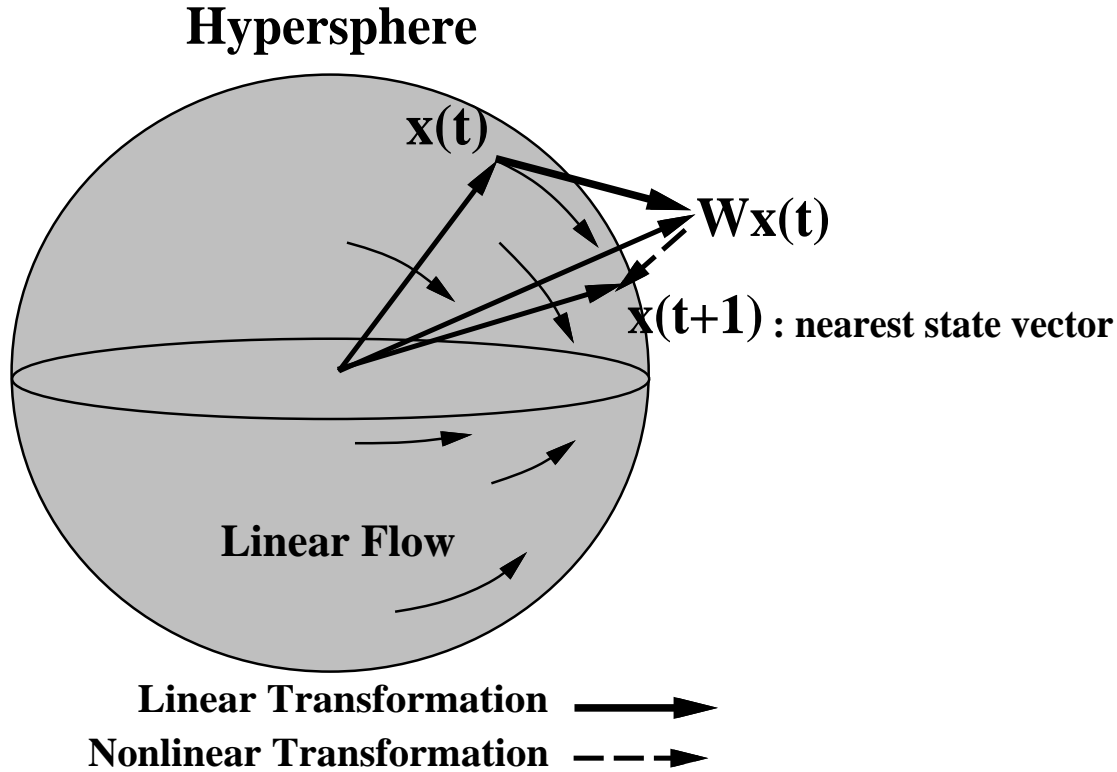


図 2.1: ニューラルネットワークのダイナミクスの幾何学的な描像。

よって線形変換される。そして、第二段階として、その変換されたベクトルは式 (2.9) で表される非線形関数によって各要素が ± 1 の格子点の中で最も近い点に移される [19]。

この見地から、ニューラルネットワークのダイナミクスを幾何的に示した図が図 2.1 である。

ここで、式 (2.8) が

$$\mathbf{x}(t+1) = \operatorname{sgn} \left(\frac{\sqrt{N}W\mathbf{x}(t)}{|W\mathbf{x}(t)|} \right) \quad (2.10)$$

と書けることに注意して、長さ \sqrt{N} の N 次元ベクトル \mathbf{v} の集合を $G_{\mathbf{v}}$ とし、ベクトル \mathbf{v} の遷移

$$\mathbf{v} \rightarrow \frac{\sqrt{N}W\mathbf{v}}{|W\mathbf{v}|} \quad (2.11)$$

を考える。ベクトル \mathbf{v} の終点は半径 \sqrt{N} の $(N-1)$ 次元超球面上にあるから、記憶行列 W によって生成されるベクトル \mathbf{v} の遷移は、超球面上の流れとして表現できる。式 (2.11) で表される線形変換を繰り返すと、ベクトル \mathbf{v} は固有値の絶対値が大きい固有ベクトル方向に進み、最終的には固有値の絶対値が最大の固有ベクトル方向に \mathbf{v} は収束する。つ

まり、式 (2.11) のフローは、超球面上において固有値の絶対値が小さい固有ベクトルで張られる空間から、固有値の絶対値が大きい固有ベクトルで張られる空間へ向かって流れている。

符号関数 $\text{sgn}(u)$ による非線形変換においては、角度変化最小で移動できる状態ベクトル (各要素が ± 1 の格子点) へ移動する。線形変換による移動量をこの非線形変換が打ち消す時、状態は遷移しなくなる。そして、ひとたび非線形変換によってもとの状態に戻されれば、以後その状態に留まる。非線形変換は、再近傍の状態ベクトルへの移動であるから、この状態遷移の停止が起こるのは、線形変換による移動量が小さい場合だけである。つまり、式 (2.8) で表される連想記憶のダイナミクスの停留点は、記憶行列で表現される線形変換によって生成される流れの淀むところに、非線形変換の安定化作用によって生み出されていると考えることができる。

ニューロンの活性化関数が連続的なシグモイド関数

$$g(u) = \frac{1 - \exp(-\beta u)}{1 + \exp(-\beta u)} \quad (2.12)$$

で与えられる場合も、超球を用いた議論を行うことができる。但し、この場合は、超球面上のフローではなく、超球の内部を含んだフローとなる。シグモイド関数を

$$g(u) = \begin{cases} 1, & \text{if } k \geq u \\ \frac{1}{k}u, & \text{if } -k \geq u < k \\ -1, & \text{if } u < -k \end{cases} \quad (2.13)$$

と近似すれば、非線形変換においては、 $-k \geq u < k$ の範囲では定数倍の線形変換が、それ以外の部分では状態が $-1, 1$ の二値に量子化されていることになる。このことから、シグモイド関数を用いた場合も、線形変換が状態遷移にドミナントに働き、非線形変換はローカルに働くという描像に変わりはない。但し、 β が大きくなる程、線形変換が保存され量子化が行われな領域 $-k \geq u < k$ が拡大するため、非線形の量子化効果は弱まり、系はより線形系に近づくことになる。

2.2.2 連続時間ダイナミクスの幾何

$(-1, 1)$ 二値出力のニューラルネットワークの連続時間ダイナミクスは、前述の通り

$$\tau \frac{du_i}{dt} = -u_i + \sum_{j=1}^N w_{ij} x_j \quad (2.14)$$

$$x_i = \text{sgn}(u_i) \quad (2.15)$$

で与えられる。この微分方程式を差分方程式に書き下すと、

$$u_i(t + \tau D) = (1 - D)u_i(t) + D \sum_{j=1}^N w_{ij} x_j(t) \quad (2.16)$$

$$x_i(t) = \text{sgn}(u_i(t)) \quad (2.17)$$

となる。ここで、 $D = 1, \tau = 1$ とおくと、前小節で述べた離散時間同期ダイナミクスとなる。以下では、連続時間ダイナミクスをこの差分方程式で時間差分 D が小さい場合に置き換えて議論を行うことにする。

連続時間ダイナミクスと離散時間ダイナミクスにおける非線形関数の働き方は全く同じである。よって、非線形変換の効果がローカルに働くという描像に変わりはない。しかしながら、線形変換が過去の内部状態の影響を残す形をしている点で、線形変換の性質が大きく異なってくる。

式 (2.16)、(2.17) で与えられる差分方程式において、符号関数による非線形変換が無いとすると、超球面上の状態遷移のフローは

$$\mathbf{u} \rightarrow \frac{\sqrt{N}((1-D)\mathbf{u} + DW\mathbf{u})}{|(1-D)\mathbf{u} + DW\mathbf{u}(t)|} \quad (2.18)$$

で与えられる。式 (2.18) で表される線形変換を繰り返すと、ベクトル \mathbf{u} は固有値の絶対値が大きい固有ベクトル方向ではなく、正の大きい固有値をもつ固有ベクトル方向に進み、最終的には固有値が最大の固有ベクトル方向に \mathbf{u} は収束する。つまり、式 (2.18) のフローは、超球面上において固有値の小さい固有ベクトルで張られる空間から、固有値の大きい固有ベクトルで張られる空間へ向かって流れている。

連続時間ダイナミクスは、以上で述べた線形フローに符号関数の量子化効果が加わったものと同じと見ることができる。但し、ここで注意すべき点は、連続時間ダイナミクスの場合、状態遷移は \mathbf{x} のみで決定されるわけではなく、 \mathbf{u} にもよるので、離散時間の時と違って、 \mathbf{x} の動きが停止しても、それはダイナミクスの収束を意味しない。連続時間ダイナミクスで系が収束した状態では

$$\mathbf{u}(t) = W\mathbf{x}(t) \quad (2.19)$$

が成立している必要がある。つまり、離散時間モデルの場合は非線形変換が線形変換を打ち消すときに状態遷移が停止するが、連続時間モデルでは式が満たされる時に状態遷移が停止する。

2.2.3 (0, 1) 出力モデルのダイナミクスの幾何

(0, 1) 出力型のステップ関数ニューロンからなるネットワークにおいては、ダイナミクス (離散時間) は

$$x(t+1) = 1\left(\sum_{i=1}^N w_i x_i(t) - h_i\right) \quad (2.20)$$

$$1(u) = \begin{cases} 1, & \text{if } u \geq 0 \\ 0, & \text{if } u < 0 \end{cases} \quad (2.21)$$

となる。この場合は、状態 \mathbf{x} は中心が $(\frac{1}{2}, \frac{1}{2}, \dots, \frac{1}{2})^T$ の半径 $\frac{\sqrt{N}}{2}$ の超球面上にあることになる。非線形変換は、この中心からみた角度変化最小点への量子化になっている。よつ

て、この場合も線形変換を中心にした議論で系のダイナミクスの概観は得られる。(0,1)出力型ネットワークの主だった特徴としては、0出力状態という結合荷重が無視される状況を含むこと、閾値の加減算を含んだ affine 変換になっていることが挙げられる。

2.3 まとめ

本章においては、ニューラルネットワークのダイナミクスの幾何構造について、状態ベクトルを張る超球面を用いた議論を行った。その結果、ニューラルネットワークのダイナミクスでは、線形変換がドミナントに働いており、非線形変換はローカルな量子化による状態の安定化を行っているという構図が得られた。本章に続く、第三章、第四章においては、本章で得られた知見をもとに、結合荷重が対称なネットワーク、非対称なネットワークの解析を行うことにする。

第 3 章： 対称ネットワークの幾何

前章では、ニューラルネットワークのダイナミクスにおいて線形変換がドミナントに働いていることを示した。このことから、荷重行列による線形変換の性質を調べること、ニューロダイナミクスの性質をある程度説明できると期待される。行列による線形変換の性質を明らかにするには、その行列の固有空間解析を行うことが有効な手立てになる。そこで、本章では、固有空間解析が容易な対称行列を荷重行列として持つ対称ネットワークのダイナミクスを幾何的に説明することを試みる。対称行列を荷重行列として持つネットワークの例として、ここでは主に自己相関連想記憶を解析対象として取り上げることにする。

3.1 自己相関連想記憶の幾何

3.1.1 自己相関連想記憶

自己相関連想記憶 [8, 17, 22, 31] では、荷重行列は

$$w_{ij} = \frac{1}{N} \sum_{\mu=1}^P s_i^\mu s_j^\mu - a\kappa\delta_{ij} \quad (3.1)$$

で与えられる。ここで、記憶パターン $\mathbf{s}^\mu = (s_1^\mu s_2^\mu \cdots s_N^\mu)^T (\mu = 1, 2, \dots, P)$ は各成分が +1、または -1 をとるベクトルであり、 δ_{ij} はクロネッカーのデルタ、 a は記憶率 $a = P/N$ である。また、 κ は自己結合の強度を決定するパラメータで、自己結合のないネットワークでは荷重行列は

$$w_{ij} = \frac{1}{N} \sum_{\mu=1}^P s_i^\mu s_j^\mu - a\delta_{ij} \quad (3.2)$$

となり ($\kappa = 1$)、自己結合を切断せずに放置する場合は荷重行列は

$$w_{ij} = \frac{1}{N} \sum_{\mu=1}^P s_i^\mu s_j^\mu \quad (3.3)$$

で与えられる ($\kappa = 0$)。この小節では、自己結合はないものとして議論を行い、自己結合がある場合については後の小節で議論することにする。

記憶の想起過程としては、全ての細胞が離散時間 $t = 0, 1, 2, \dots$ に同期して状態遷移するモデルを考える。時刻 t における状態を $\mathbf{x}(t) = (x_1(t)x_2(t)\cdots x_N(t))^T$ とすると、そのダイナミクスは

$$x_i(t+1) = \text{sgn}(u_i(t)) \quad (3.4)$$

$$\text{sgn}(u) = \begin{cases} 1, & \text{if } u \geq 0 \\ -1, & \text{if } u < 0 \end{cases} \quad (3.5)$$

で与えられる。

ここでは、ランダムパターンを記憶パターンとして相関学習を行う場合について考える。この系については既に多くの研究がなされており、その性質が詳細にわたり明らかにされている。本小節では、現象論として知られている主だった特徴を以下に列挙することにする。

まず、記憶パターンは引き込み領域をもつ系の安定点になる。よって、ノイズを含むパターンから記憶パターンを想起することができる。図 3.1 に想起ダイナミクスにおける記憶パターン \mathbf{s} と状態ベクトル \mathbf{x} のオーバーラップ (方向余弦)

$$m(t) = \frac{1}{N} \mathbf{s}^T \cdot \mathbf{x}(t) \quad (3.6)$$

の変化をプロットした図を示す。この図のように、初期状態が記憶パターンに近い時は記憶パターンがダイナミクスの収束点となるが、ある程度以上の距離があると他の状態へと引き込まれる。この臨界距離で引き込み領域は定義される。

また、ネットワークには記憶容量が存在し、記憶するパターン数が増えると、ある時点で系が連想記憶として機能しなくなる。すなわち、記憶パターンが引き込み領域を持つ安定点ではなくなる。図 3.2 に記憶容量以上にパターンを記憶した場合の想起ダイナミクスにおけるオーバーラップの変化を示す。この様に、記憶容量を超えると記憶パターン付近のパターンも記憶パターンから離れ、記憶パターンベクトルと異なる状態ベクトルに引き込まれてしまう。また、図 3.3 に、各記憶率での引き込み領域と記憶パターンの安定性 (記憶パターンを初期値としたときの収束先のパターンと記憶パターンとのオーバーラップ) を示す。一般に、記憶容量は引き込み領域の曲線と安定性の曲線が交わる時の記憶率で定義する。この図で見ると、記憶率 $a = 0.15$ 付近が記憶容量となっているが、これはよく知られた自己相関連想記憶の性質である。

次の小節では、以上で述べた自己相関連想記憶の性質を、前節で述べた幾何学的な視点から考察する。

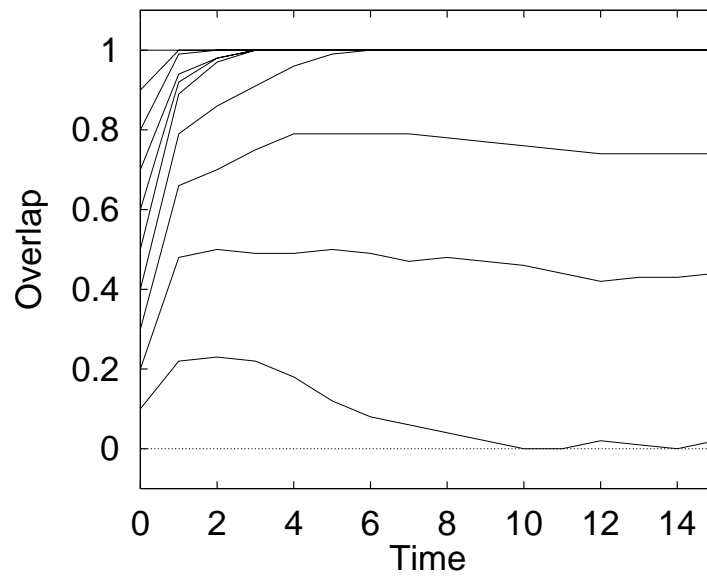


図 3.1: 自己相関連想記憶の想起過程におけるオーバーラップの変化 (自己結合なし, $a = 0.1, N = 500$)。

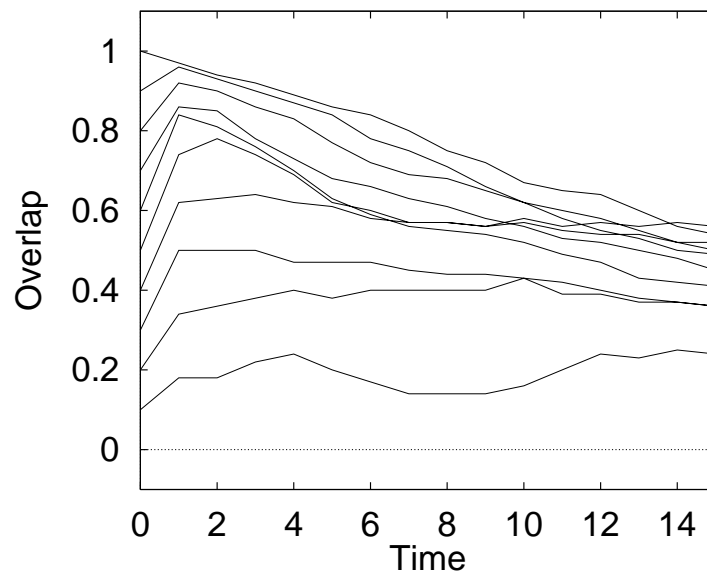


図 3.2: 自己相関連想記憶の想起過程におけるオーバーラップの変化 (自己結合なし, $a = 0.1, N = 500$)。

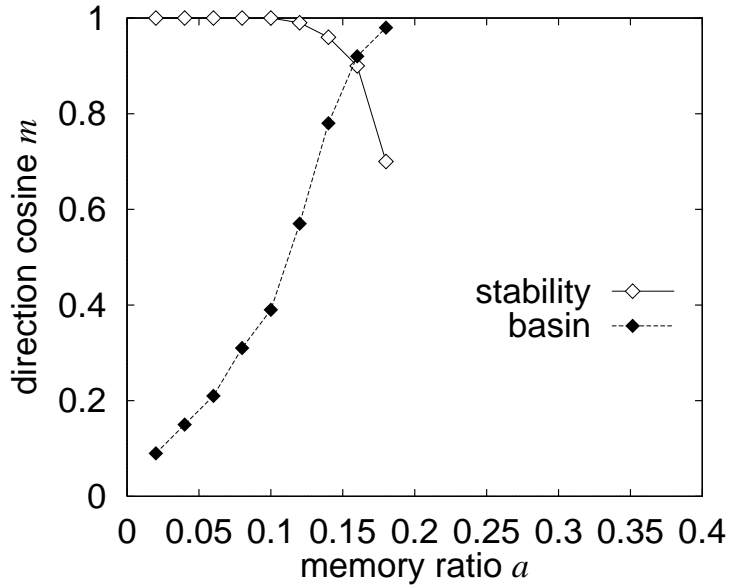


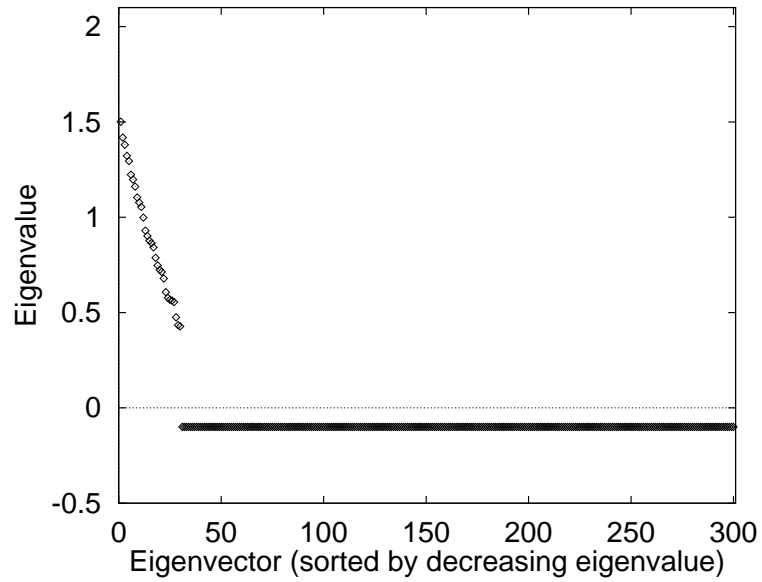
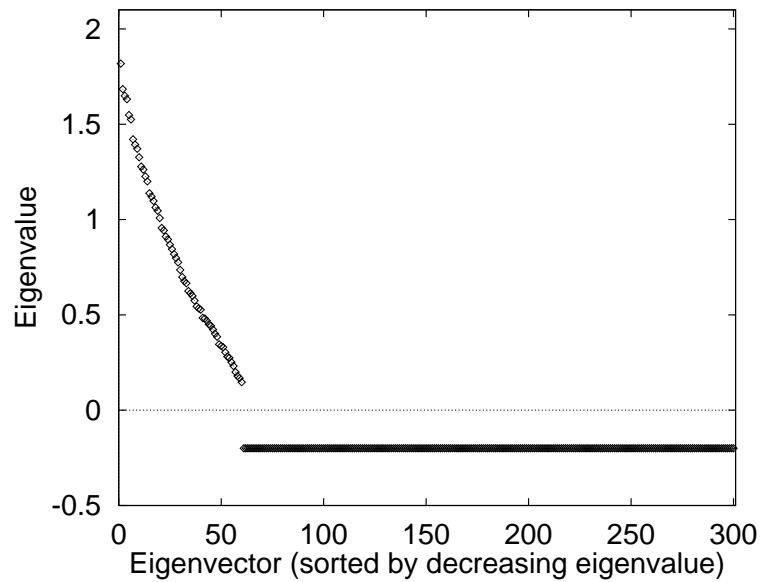
図 3.3: 自己相関連想記憶 (自己結合なし) における記憶パターンの安定性と引き込み領域 ($N = 500$)。

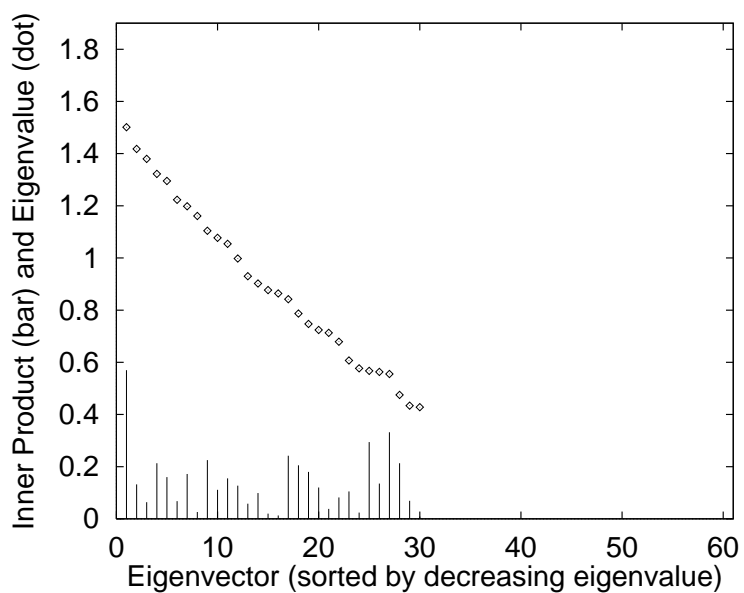
3.1.2 自己相関連想記憶の幾何

第2章で述べたように、ニューロダイナミクスにおいては荷重行列による線形変換が系の状態遷移の原動力となっている。よって、荷重行列の固有空間の構造を調べることで、系の力学的性質が明らかになると期待される。そこで、まず、 $N = 300$ 個のニューロンからなるネットワークに $P = 30$ 個のパターン ($a = 0.1$) と $P = 60$ 個のパターン ($a = 0.2$) を記憶した場合について、固有値の分布 ($\kappa = 1$) を調べることにする (図 3.4)。図 3.4 から分かるように、固有空間は P 個の正の固有値と、 $(N - P)$ 個の $-a$ の固有値からなっている。そして、正の固有値の最大、最小の幅は記憶パターン数 P が増加するにしたがって大きくなる。

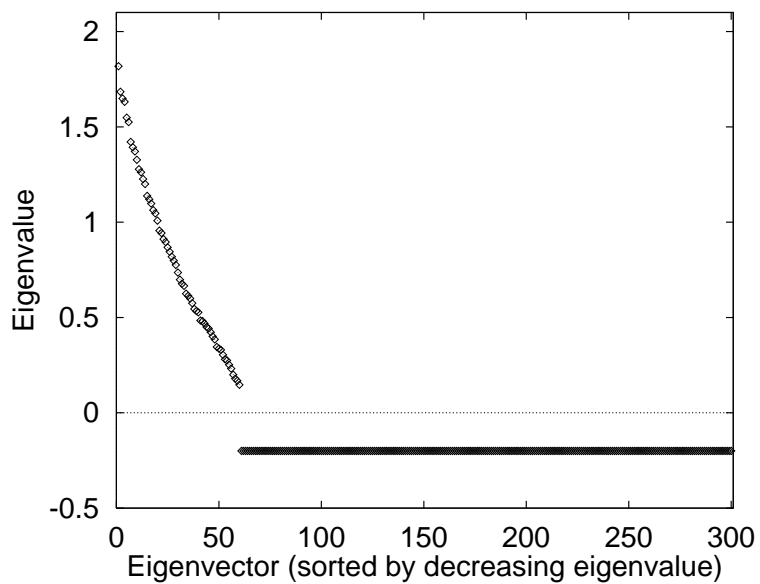
次に、記憶パターンと固有ベクトルの関係を調べる。図 3.5 に $N = 300, P = 30, 60$ の場合の記憶パターンの一つと固有ベクトルの内積を示す。この図のように、記憶パターンは正の固有値を持つ固有ベクトル全てに分散して表現されており、それ以外の空間とは直交している。つまり正の固有値をもつ固有ベクトルが記憶パターンを含む空間 (記憶空間) を張り、それ以外の固有ベクトルが記憶と直交する空間 (ノイズ空間) を形成している。この事実と、固有値の分布を合わせて考えると、記憶パターンへの引き込み、及び記憶容量の存在は次のように説明できる。

まず、記憶パターンへの引き込みであるが、荷重行列は記憶パターンと直交する空間では固有値は $-a$ で、記憶パターン数が少ない間は小さい値に保たれる。よって、線形

(a) W の固有値の分布 ($a = 0.1$)(b) W の固有値の分布 ($a = 0.2$)図 3.4: 自己相関連想記憶行列 W の固有値の分布 ($N = 300, \kappa = 1$)。



(a) 固有ベクトルと記憶パターンの内積 ($a = 0.1$)



(b) 固有ベクトルと記憶パターンの内積 ($a = 0.2$)

図 3.5: 自己相関連想記憶行列 W の固有ベクトルと記憶パターンの内積 ($N = 300, \kappa = 1$)。

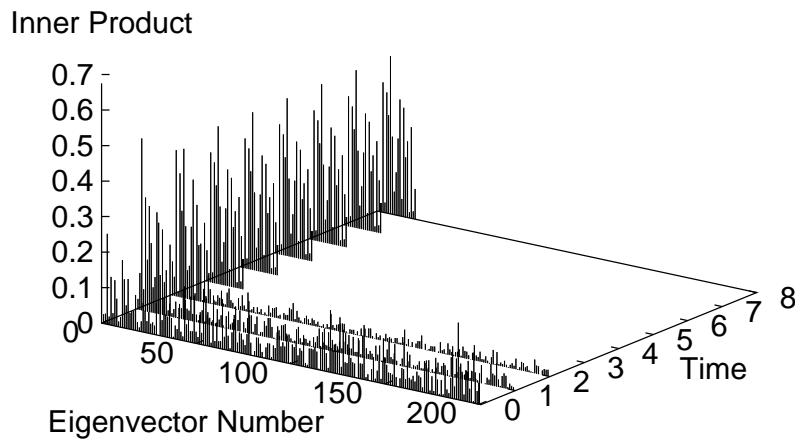


図 3.6: 想起過程における固有ベクトル成分の変化 ($a = 0.1, N = 200$)。初期状態と記憶パターンのオーバーラップは 0.4。ここで、固有ベクトル番号 (Eigenvalue Number) は固有値の大きい順である。

変換によってノイズ空間の固有ベクトル成分は減少する。もし、符号関数による量子化がなければ、系は絶対値最大の固有値を持つ固有ベクトル方向に引き込まれる。しかし、符号関数の量子化により、記憶空間の固有値に大きな開きが無い場合は記憶空間に入ると同時に状態遷移は停止する。よって、記憶率が低く、記憶空間内の最大固有値と最小固有値の差が小さい間は、記憶空間は安定となり、記憶空間内の全固有ベクトル成分をもつ記憶パターンは系の安定点となる。この説明を裏付けるため、想起過程における各固有ベクトル成分の増減を調べた結果を図 3.6 に示す。想起過程において、ノイズ空間成分が消滅し、記憶空間成分が増加した結果、記憶パターンが想起される様子がこの図によく示されている。

記憶率が高くなり、記憶空間内の固有値の分散が増大すると、記憶空間内でも線形変換による角度変化が増し、符号関数の安定化効果を超えてしまう。その結果、記憶パターンが不安定になり、記憶の想起ができなくなる。これが記憶容量を超えた時、記憶パターンが不安定になる幾何的メカニズムである。この時生じる偽記憶パターンについて次の小節で扱うことにする。

3.1.3 偽記憶

この小節では、記憶パターンが不安定になったときに出現する偽記憶について検討する。前の小節で述べたとおり、記憶率が記憶容量を超え、記憶パターンの想起に失敗したとき、系のダイナミクスの収束先として偽記憶パターンが現れる。これを幾何的に説明すると以下ようになる。

前小節で、記憶率が高い場合、記憶空間内の固有値の分散が増大することで記憶空間内でも線形変換による角度変化が増し、記憶パターンが不安定になると主張した。その時、この系は絶対値最大の固有値を持つ固有ベクトル方向へ向かう。つまり、偽記憶パターンとは絶対値最大の固有値を持つ固有ベクトルに近いパターンであると予想される。実際、偽記憶パターンの一つを固有ベクトルで展開すると、最大固有値をもつ固有ベクトル成分を大きく持つ(図 3.7)。また、系が偽記憶に引き込まれるときの各固有ベクトル成分の増減を図 3.8 に示す。この図から、想起過程において、大きい固有値を持つ固有ベクトル成分が徐々に増加し、最終的に最も大きい固有値を持つ固有ベクトル群で張られる偽記憶状態へ系が収束していく様子を見ることができる。

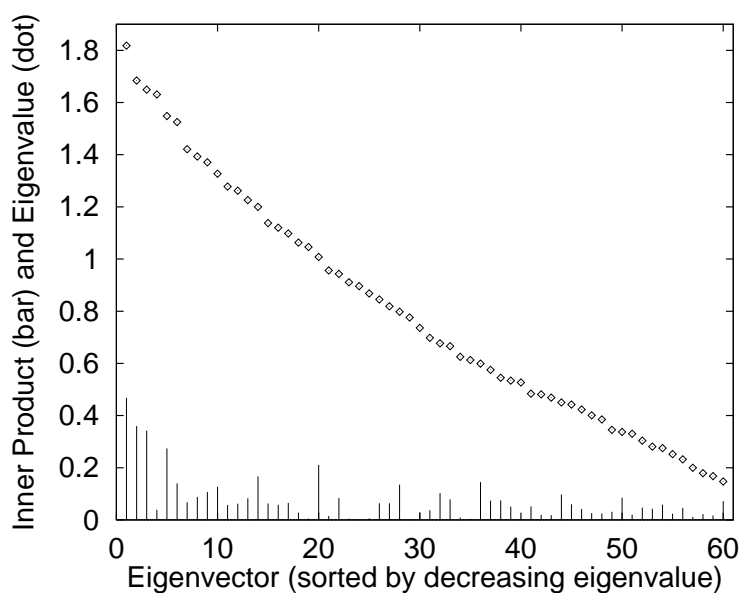


図 3.7: 自己相関連想行列 W の固有ベクトルと偽記憶パターンの内積 ($a = 0.2, N = 300$)。

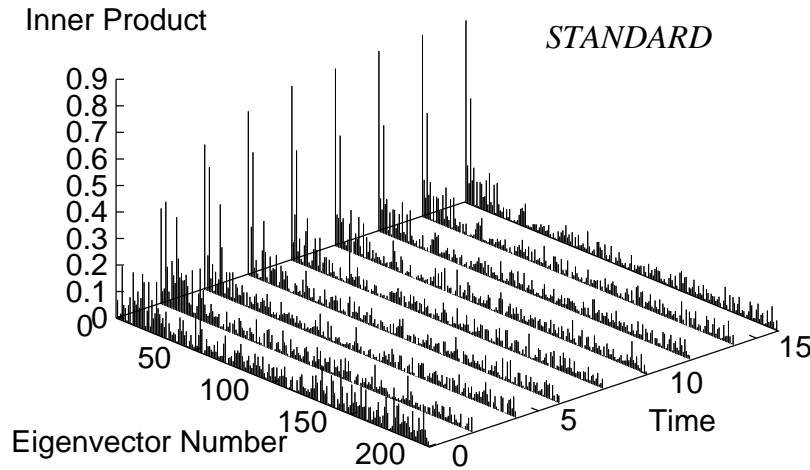


図 3.8: 記憶容量超過時の想起過程における固有ベクトル成分の変化 ($a = 0.2, N = 200$)。

3.1.4 部分反転法

自己相関型連想記憶の記憶容量を向上させるダイナミクスとして、部分反転法が Morita によって提案されている [29]。部分反転法とは、

$$x_i(t+1) = \text{sgn}\left(\sum_{j=1}^N w_{ij}(x_j - \lambda\phi(u_j(t)))\right) \quad (3.7)$$

という二段階ダイナミクスを使用する方法である。ここで、

$$u_j(t) = \sum_{k=1}^N w_{jk}x_k(t) \quad (3.8)$$

で、また、関数 $\phi(u)$ は

$$\phi(u) = \begin{cases} -1, & \text{if } u < -h \\ 0, & \text{if } -h \leq u < h \\ 1, & \text{if } h \leq u \end{cases} \quad (3.9)$$

で与える (図 3.9)。

この想起ダイナミクスを用いると、記憶容量が $0.25N$ から $0.30N$ 程度まで上昇することが知られている。部分反転法を用いたときの想起過程におけるオーバーラップの変化を図 3.10($a = 0.2$)、図 3.11($a = 0.4$) に示す。この図から分かる通り、部分反転法による

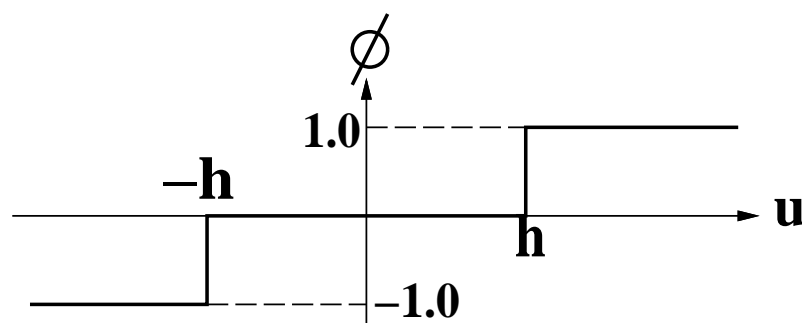


図 3.9: 部分反転法に用いる関数。

想起の特徴として、記憶容量が向上することに加え、偽記憶が振動状態となる現象が見られる。この部分反転法による記憶容量の向上と偽記憶の不安定化という現象を幾何的に捉えると以下のような説明が成り立つ。

部分反転法は、内部状態 u_i の絶対値が大きいニューロンの影響を打ち消し、そのような u_i を含む状態を不安定化する。連続時間ダイナミクスにおいては非単調ニューロンを用いることによって同様の効果が得られる。これらの現象の理論解析は Shiino[41] や Yoshizawa[44] らによって明らかにされているが、ここでは連想記憶の幾何学的な解析法の立場からその原理をより簡潔に説明する。

絶対値の大きい u_i を含む状態は、固有値の大きい固有ベクトルに近いパターンに対応している。前小節で述べたように、これが偽記憶パターンに対応している。よって、部分反転法は偽記憶を不安定化していることに対応する。実際、部分反転法を用いた想起過程における固有ベクトル成分の変化を図 3.12、図 3.13 に示す。図 3.12 のように、記憶率が 0.2 では、大きい固有値を持つ固有ベクトル成分の増加が抑えられ、想起が実現している。図 3.13 のように、記憶率が 0.4 となると、記憶パターンの想起はできないが、大きい固有値をもつ固有ベクトル方向への流れを引き戻す力が働き、その空間成分の増加が抑えながら振動状態に突入している様子が見てとれる。以上の現象を超球面を用いて図示すると、図 3.14 のような描像で系を捉えることが可能である。

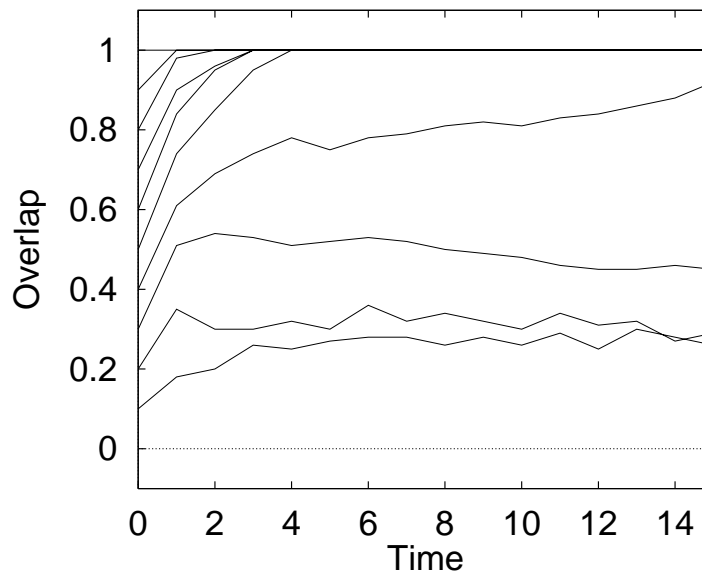


図 3.10: 部分反転法を使って想起に成功する場合 (記憶率 $a = 0.2$) のオーバーラップの変化 (自己結合なし, $N = 500, h = 1.5, \lambda = 1.0$)。

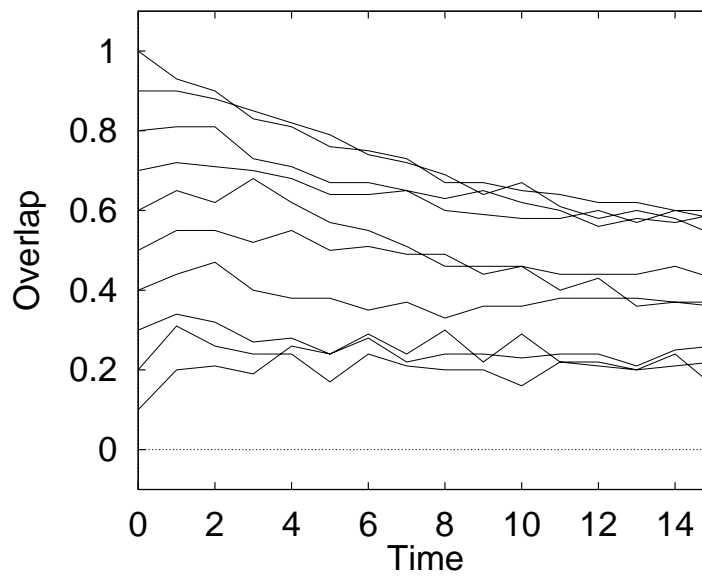


図 3.11: 部分反転法を使っても想起が不可能な場合 (記憶率 $a = 0.4$) のオーバーラップの変化 (自己結合なし, $N = 500, h = 1.5, \lambda = 1.0$)。

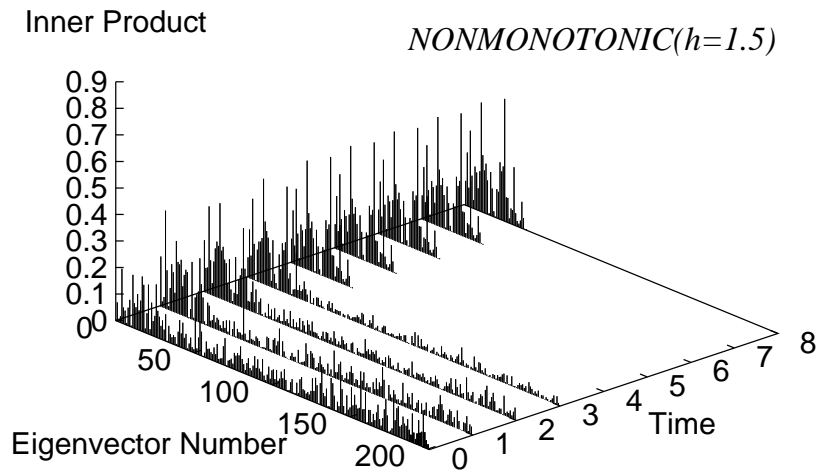


図 3.12: 部分反転法による想起過程での固有ベクトル成分の変化 (想起成功時, $a = 0.2, N = 200, h = 1.5, \lambda = 1.0$)。

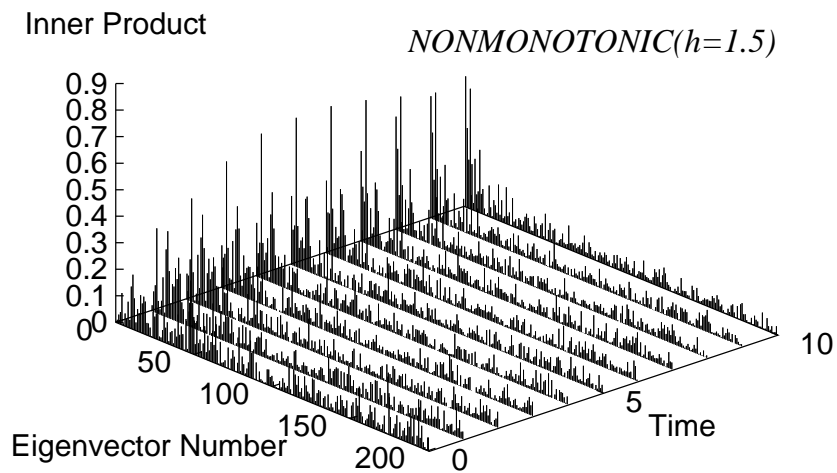


図 3.13: 部分反転法による想起過程での固有ベクトル成分の変化 (想起失敗時, $a = 0.4, N = 200, h = 1.5, \lambda = 1.0$)。

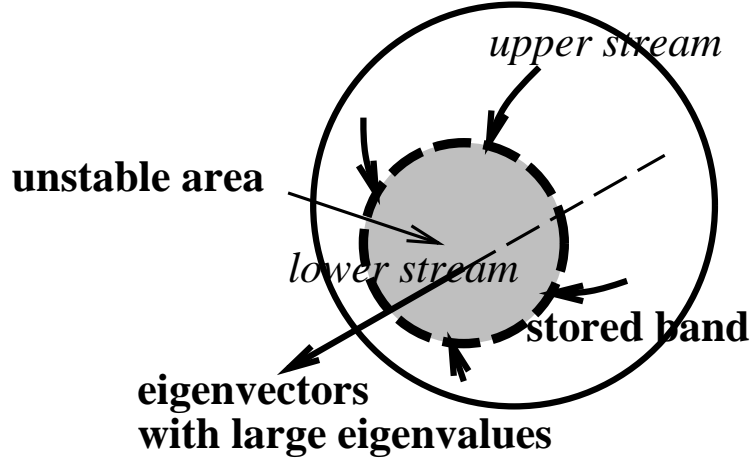


図 3.14: 部分反転法の幾何的な説明。部分反転法は内部状態の絶対値が大きくなるニューロンの影響を除去することで絶対値の大きい固有値をもつ空間への流れを塞ぎ止めている。その結果、記憶パターンを含む領域 (stored band) で状態遷移が停止し、記憶パターンの想起が可能になる。

3.1.5 想起過程の第1ステップの解析

この小節では、想起過程の第一ステップに焦点を絞り、その振舞を幾何的な視点から調べることにする。想起過程の第一ステップの特徴として、想起に成功する場合も偽記憶に収束する場合も、系のオーバーラップは大きくなる点が挙げられる。本小節では、その現象を幾何的な視点から定量的な解析も含めて論じることにする。

今、記憶パターン s^μ の想起過程を考える。初期状態 $\mathbf{x}^\mu(0)$ を、記憶パターン成分 $\mathbf{m}_\mu(0)$ 、記憶パターン以外の記憶空間成分 $\overline{\mathbf{m}}_\mu(0)$ 、ノイズ空間成分 $\mathbf{n}(0)$ に分けると、

$$\mathbf{x}^\mu(0) = \mathbf{m}_\mu(0) + \overline{\mathbf{m}}_\mu(0) + \mathbf{n}(0) \quad (3.10)$$

$$|\mathbf{m}_\mu(0)|^2 + |\overline{\mathbf{m}}_\mu(0)|^2 + |\mathbf{n}(0)|^2 = N \quad (3.11)$$

と書ける。オーバーラップ $\cos \theta_\mu = |\mathbf{m}_\mu|/\sqrt{N}$ を固定し、その他の成分をランダムに与えたとすると、記憶パターンが張る記憶空間の次元は P であり、ノイズ空間の次元は $(N - P)$ であるから、

$$|\overline{\mathbf{m}}_\mu(0)| \simeq \sqrt{N - |\mathbf{m}_\mu(0)|^2} \sqrt{\frac{P-1}{N}} \quad (3.12)$$

$$|\mathbf{n}(0)| \simeq \sqrt{N - |\mathbf{m}_\mu(0)|^2} \sqrt{\frac{N-P}{N}} \quad (3.13)$$

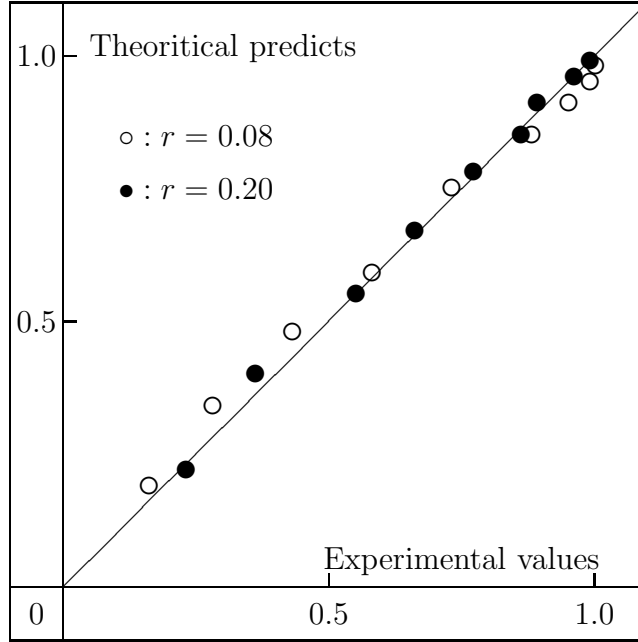


図 3.15: 1ステップ後のオーバーラップの理論値と実験値の比較。

となる。これを自己結合のない記憶行列で変換すると、正の固有値の平均が $(N - P)/N$ であり、負の固有値は $-P/N$ であるから、 $W^0 \mathbf{x}(0)$ の3つの成分の大きさはそれぞれ、

$$|\mathbf{m}_\mu(W^0 \mathbf{x}^\mu(0))| \simeq \frac{N - P}{N} |\mathbf{m}_\mu(0)| \quad (3.14)$$

$$|\overline{\mathbf{m}}_\mu(W^0 \mathbf{x}^\mu(0))| \simeq \frac{N - P}{N} |\overline{\mathbf{m}}_\mu(0)| \quad (3.15)$$

$$|\mathbf{n}(W^0 \mathbf{x}^\mu(0))| \simeq \frac{P}{N} |\mathbf{n}(0)| \quad (3.16)$$

となる。

非線形変換が最近接の格子点を選ぶだけであることを考えれば、線形変換につづく非線形変換によってもこの成分の比は大きく変化しないと仮定できる。よって、 $\mathbf{x}^\mu(1)$ のオーバーラップは、記憶率が a のとき、式 (3.10)(3.12)(3.13)(3.14)(3.15)(3.16) より、

$$\cos \theta_\mu(1) \simeq \frac{|\mathbf{m}_\mu(0)|}{\sqrt{|\mathbf{m}_\mu(0)|^2 + (N - |\mathbf{m}_\mu(0)|)^2 a(1 - a)^{-1}}}$$

となる。ここで求めた理論値と、実験値との比較をしてみると、図 3.15 のようにこの理論値と実験値がよく一致することがわかる。

以上の議論から、想起過程の第一ステップでオーバーラップが増加するのは、記憶パターンから離れた状態ではノイズ空間成分を持っているため、荷重行列による線形変換でノイズ空間成分が減少し記憶空間成分が増加することによるという説明が得られた。

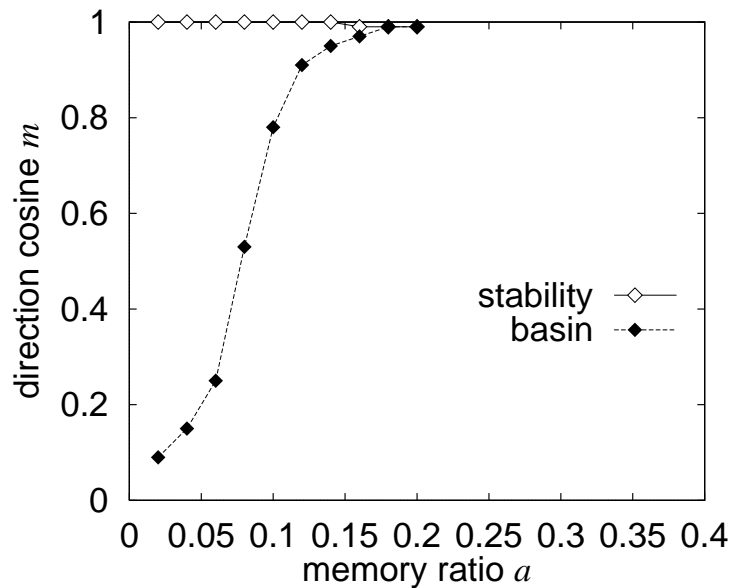


図 3.16: 自己相関連想記憶 (自己結合あり) における記憶パターンの安定性と引き込み領域 ($N = 500$)。

3.1.6 自己結合の効果

本節のこれまでの議論は、全て自己結合がないことを仮定していた。この小節では、ネットワークが自己結合をもつ場合、自己相関連想記憶にどのような性質の違いが生じるかを論じ、固有空間解析をもとにした幾何学的な視点を導入して考察を行う。

前述の通り、連想記憶の性能としては記憶パターンの安定性、記憶パターンの引き込み領域という二つの評価基準があげられる。ネットワークが自己結合を持つ場合、各記憶率においてのこれらの値を調べた結果を図3.16に示す。自己結合がない場合と比較すると、パターンの安定性は増すが、引き込み領域は狭くなっていることが分かる。図3.16では、安定性と引き込み領域の曲線の交差点は $a = 0.18$ 程度であり、自己結合の無いモデルより記憶容量が高く出ているが、 N の値が大きくなると記憶容量は自己結合の無いモデルに近づくことが分かっている [12]。

自己結合を導入すると系全体の可動性が減少する。そのため、記憶の安定性は増すが、同時に記憶以外の状態も安定化され、その結果引き込み領域が小さくなると考えられる。図3.17、3.18に、自己結合を導入した場合の想起過程のオーバーラップの変化を示す。系の可動性の現象がこの図からもよく見てとることができる。この可動性の低下に幾何学的視点から解釈を加えると以下のような説明することができる。

自己結合のある場合の固有値の分布を計算すると図3.19のようになる。この分布を自

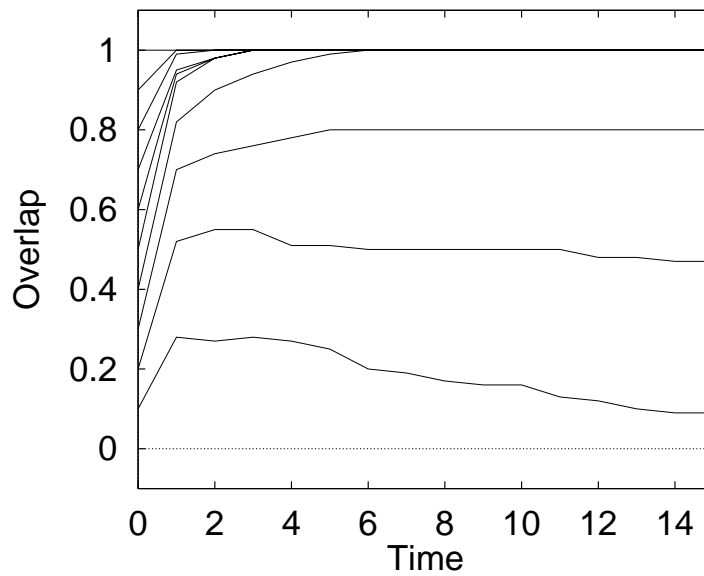


図 3.17: 自己相関連想記憶の想起過程におけるオーバーラップの変化 (自己結合あり, $a = 0.1, N = 500$)。

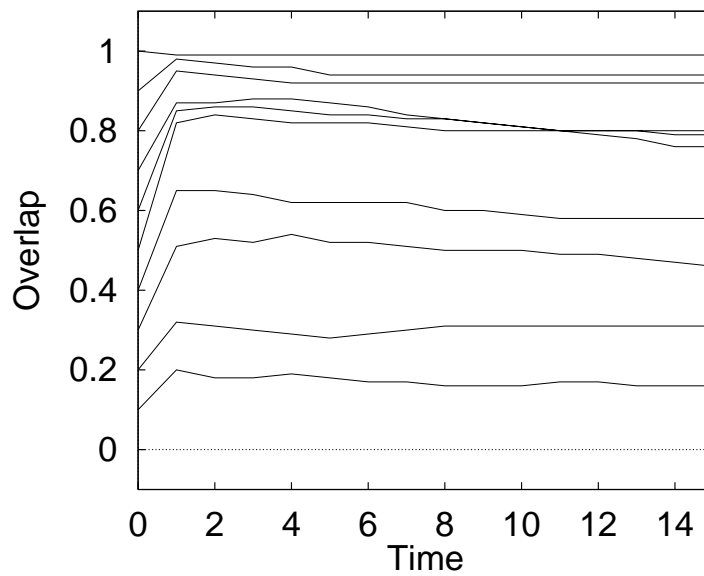
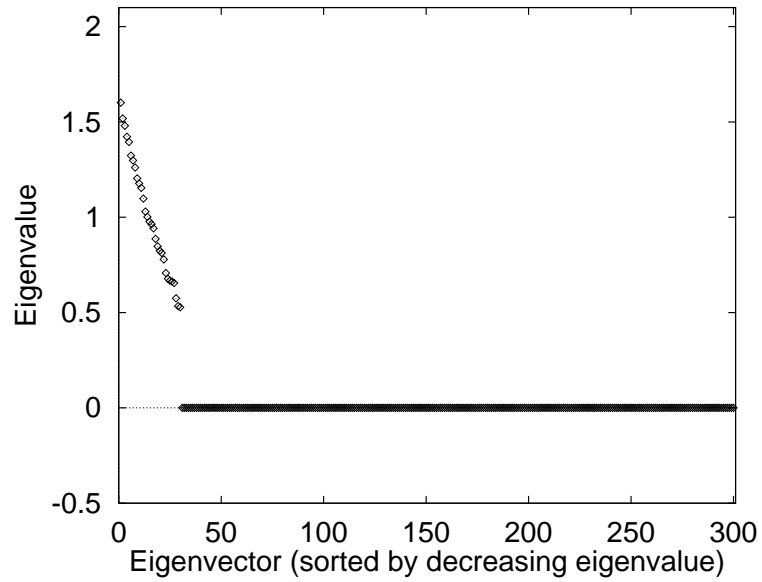
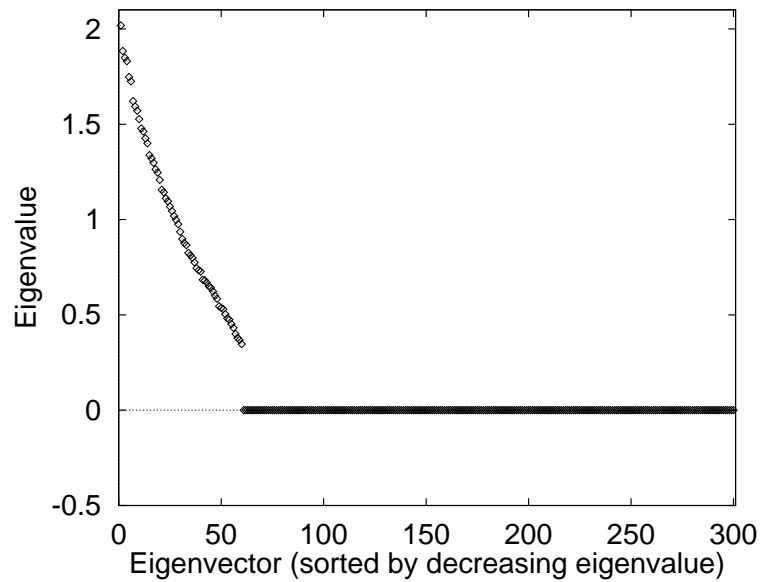


図 3.18: 自己相関連想記憶の想起過程におけるオーバーラップの変化 (自己結合あり, $a = 0.2, N = 500$)。



(a) W の固有値の分布 ($a = 0.1$)



(b) W の固有値の分布 ($a = 0.2$)

図 3.19: 記憶行列 W の固有値の分布 ($N = 300, \kappa = 0$)。

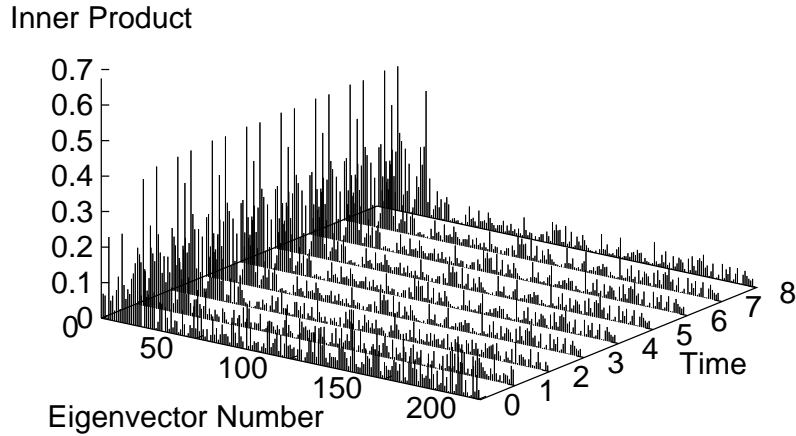


図 3.20: 記憶容量超過時の想起過程における固有ベクトル成分の変化 (自己結合あり, $a = 0.2$, $N = 200$)。初期状態と記憶パターンのオーバーラップは 0.6。

自己結合がない場合(図 3.4)と比較すると、自己結合を含む場合は全ての固有値が a だけ大きくなっていることが分かる。自己結合の導入は、荷重行列に対角行列を加えることに対応することを考慮すれば、行列 $aI(I$: 単位行列) を足せば、固有値が a 上昇し、固有ベクトルは変化しないことは自明である。

前章で述べた通り、線形変換による角度変化が大きい場合は系の可動性は高くなる。線形変換による角度変化は状態を張る固有ベクトルの固有値の分散が増すほど大きくなる。但し、線形変換による角度変化を決定するのは固有値の差ではなく固有値の比になる。全固有値を大きくすれば、記憶空間内の固有値の比は小さくなる。これが記憶空間付近の状態ベクトルの可動性を下げていると考えられる。図 3.20 に自己結合を含むネットワークで偽記憶に引き込まれる時の固有ベクトル成分の変化を示す。この図から、自己結合を含む場合、大きい固有値を持つ固有ベクトル成分の増加によって偽記憶状態に陥るのではなく、記憶空間への引き込みが進まない結果、記憶が想起できないことが分かる。

自己結合を含むモデルで、記憶パターン数が増すことによって引き込み領域が消滅することに対して、固有値の分布から原因を考えると、記憶パターン数が増加することによりノイズ空間の次元が減少するという要因が挙げられる。しかし、第9章で後述する

が、直交学習法により固有値分布の広がりを抑制する場合は、ノイズ空間の次元が高くなっても引き込み領域をある程度確保することができる。このことから、自己結合による分布の広がりの抑制は、自己結合によらない固有値分布の抑制とは異なる意味を持つと考えられる。この点を理解するには非線形の量子化効果と自己結合の関係を定量的に議論する必要があり、本章の枠組みを超えた議論が必要になる。非線形の効果も含めた定量的議論は第6章に譲ることとする。

3.1.7 D関数

以上の議論では、ニューラルネットワークのダイナミクスを固有空間解析を用いて説明した。しかし、ニューロン数が増えると、全ての固有ベクトルとの内積をモニターするのは困難になる。そこで、本節では系のダイナミクスの動きをD関数 [10, 13] という一次元の量を導入することで解析を進めることにする。

連想記憶においては、記憶パターンとのオーバーラップを指標としてダイナミクスを捉えるのが一般的である。幾何的な解析においては、ダイナミクスを固有ベクトル成分の動きとして捉えた。ここでは、系のダイナミクスの幾何的な動きの指標として金道、筆者によって導入されたD関数について説明する。

今まで述べてきたように、幾何的な視点からダイナミクスを見ると、荷重行列の固有ベクトルのうち、固有値の絶対値が大きいものがその成分を増やしていく。であるから、その動きをモニターする一次元量を導入できれば、それを指標として状態遷移の様子を容易に観測できる。非同期型確率的状態遷移、及び連続時間ダイナミクスの場合は、関数

$$E(\mathbf{x}(t)) = -\frac{1}{2}\mathbf{x}(t)^T W \mathbf{x}(t) \quad (3.17)$$

がエネルギー関数になっている。しかし、離散時間同期型ダイナミクスの場合は、上記の関数はエネルギー関数とはならない。離散時間同期型ダイナミクスでは、関数

$$E_d(\mathbf{x}(t)) = -\frac{1}{2}\mathbf{x}(t+1)^T W \mathbf{x}(t) \quad (3.18)$$

$$x_i(t+1) = f\left(\sum_{j=1}^N w_{ij}x_j(t)\right) \quad (3.19)$$

がエネルギー関数となる(付録1 参照) [20]。但し、ここでニューロンの活性化関数 $f(u)$ は単調増加であることを仮定している。

離散時間同期型ダイナミクスのエネルギー関数を状態ベクトル $\mathbf{x}(t)$ と荷重行列 W だけで定義するには、非線形関数 $f(u)$ が含まれるので計算が複雑になる。そこで、この非線形関数を省いた関数

$$D(\mathbf{x}) = \frac{1}{N}\mathbf{x}^T W^T W \mathbf{x} \quad (3.20)$$

を考えることができる。これがD関数と定義される。ここで、D関数はエネルギーと異なり、負の係数が付加されていないことに注意を要する。

このD関数は

$$D(\mathbf{x}) = \frac{1}{N} \sum_{i=1}^N \lambda_i^2 C_i^2(\mathbf{x}) \quad (3.21)$$

と書き直すことができる。ここで

$$C_i(\mathbf{x}) = \mathbf{e}_i^T \mathbf{x} \quad (3.22)$$

である。このことから、D関数は絶対値の大きい固有値をもつ固有ベクトル成分を多く含むとその値が大きくなる関数であることが分かる。前述の通り、連想記憶の想起過程においては、絶対値の大きい固有値を持つ固有ベクトル成分が増加するので、D関数は時間を追って増加する。

自己結合がない記憶行列の場合は、記憶率 a の時、 $u_i^\mu = \sum_j w_{ij} s_j^\mu$ は十分大きい N に対して、平均1、分散 a になることが知られているから、 $D(\mathbf{s}^\mu)$ の期待値 D_m は $D_m = (1+a)$ となる。自己結合がある記憶行列の場合は、平均が $(1+a)$ になるので、 $D(\mathbf{s}^\mu)$ の期待値 D_m は、 $D_s(1+3a+a^2)$ になる。この平均と分散は記憶パターンに依存しないから、記憶パターン \mathbf{s} は $(N-1)$ 次元球面 $S_{\sqrt{N}}^{N-1}$ と超楕円球面

$$\frac{1}{N} \mathbf{x}^T W^T W \mathbf{x} = D_m \quad (3.23)$$

の交差面に存在していることになる (図 3.21)。この交差面のことを記憶バンド (stored band) と呼ぶ。状態ベクトルを張る $(N-1)$ 次元球面 $S_{\sqrt{N}}^{N-1}$ は、記憶バンドによって $D(\mathbf{x}) < D_m$ の部分と $D(\mathbf{x}) > D_m$ の部分に分割されることになる。

自己相関連想記憶で記憶率がそれ程大きくない間は、ノイズ空間の次元が圧倒的に大きいため、パターンは一般にノイズ成分を多くもつことになる。ノイズ空間の固有値は $-a$ であるから、記憶率が低い場合は小さい。その場合、パターンは一般に小さい固有値をもつ固有ベクトル成分を多く持つため、D関数は小さい値をとる。そして、状態遷移が進むにつれ、線形変換によって大きい固有値を持つ固有ベクトルの方向に状態ベクトルが移動すると、D関数の値は大きくなる。つまり、想起ダイナミクスにおいては、D関数の小さい領域から大きい領域への流れが生じていることになる。そこで、分割された領域のうち、 $D(\mathbf{x}) < D_m$ の領域を上流、 $D(\mathbf{x}) > D_m$ の領域を下流と呼ぶことにする。

自己相関連想記憶の幾何的な解析として、上の議論では自己結合の無い場合、自己結合のある場合、部分反転法を用いた場合のダイナミクスを論じてきた。そこで、これらのダイナミクスにおいてD関数がどのような振る舞いを示すかを以下に調べる。

まず、自己結合が無い場合についてD関数の振る舞いを調べる。記憶率が記憶容量以下の時は、状態ベクトルの遷移は記憶バンドで終了し、記憶したパターンが想起される (図 3.22)。ここで、特徴的なのは、D関数が一度 D_m を越えてから戻ってくる現象がみられることである。これは、記憶パターン付近で非線形関数 $\text{sgn}(u)$ による状態の引き戻しが起こっていることによると考えられる。この現象は固有ベクトル成分の変化を調べた図 3.6にも現れている。一方、記憶率が記憶容量を超過した場合は、D関数の値は D_m を遥かに超えていく (図 3.23)。図 3.8で、偽記憶状態は大きい固有値をもつ固有ベクトル

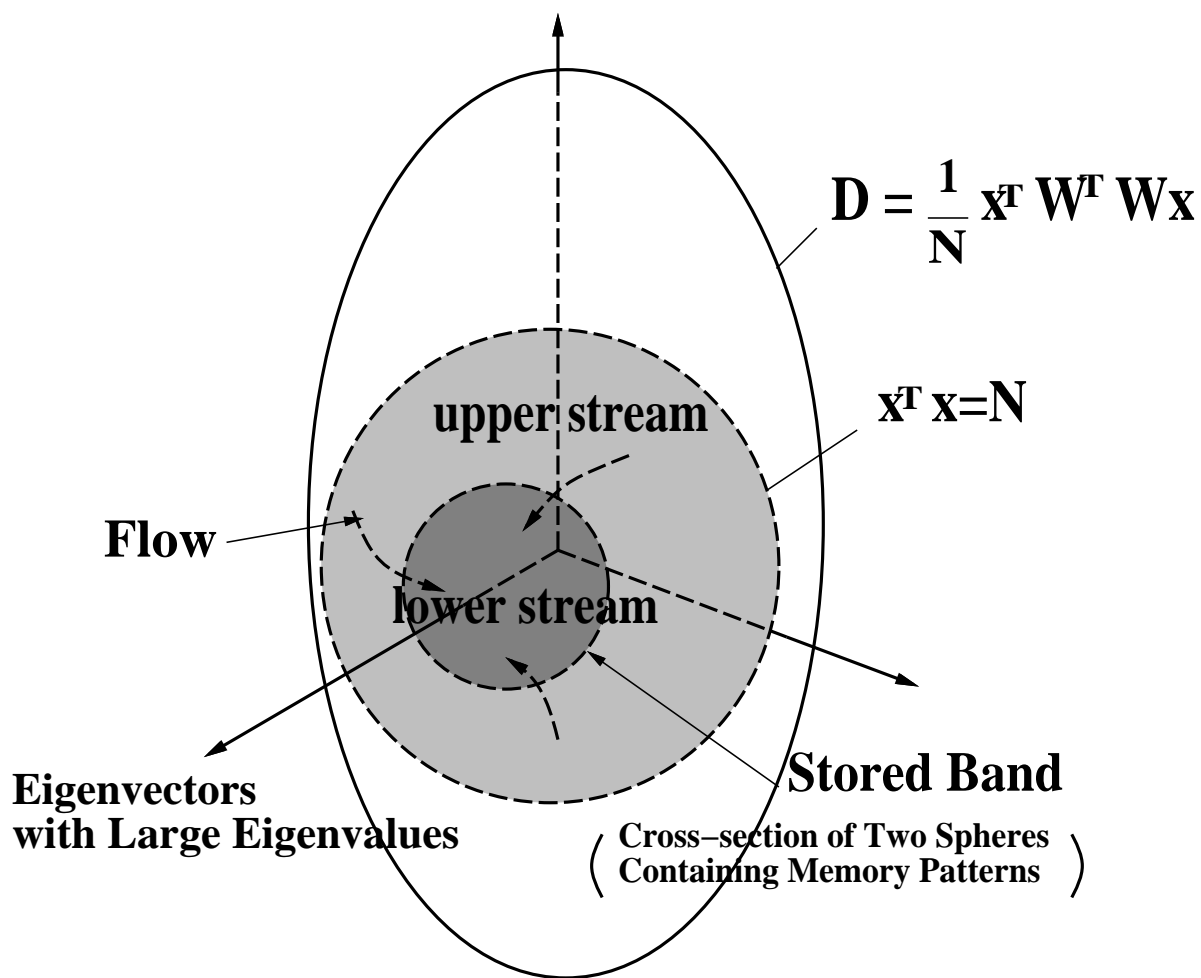


図 3.21: D 関数の幾何学的描像。

ル成分の増大によって生じる様子を示したが、このことがD関数の増加という形でもモニターできることが分かる。

次に、自己結合がある場合についてD関数の振る舞いを調べる。記憶率が記憶容量以下の時は、自己結合がない場合と大きな差はない。但し、想起に成功しない時の収束先が自己結合のない場合に比べて記憶パターンに近い傾向がある(図3.24)。記憶率が記憶容量を越えた場合は、自己結合がない場合とはD関数の振舞が大きく異なる。自己結合がある場合は、記憶が失われて偽記憶状態に収束する場合でも、D関数はそれ程増加しない(図3.25)。これは、図3.20に示した状態ベクトルの固有空間解析で得た結果をよく反映している。

最後に、部分反転法を用いた場合(自己結合無し)についてD関数の振舞いを調べる。記憶率が部分反転法の記憶容量以下の場合は、記憶バンドに状態ベクトルが引き込まれる。記憶率が部分反転法の記憶容量を超えた場合は、D関数の増加は偽記憶へ収束する時ほど大幅ではなく、記憶バンド付近で振動状態に陥る。ここで観測されたD関数の振る舞いも、固有空間解析の特徴をよく捉えていると言える。

以上で述べたように、個々の固有ベクトル成分まで細かく分析をしなくても、D関数を用いることで、固有空間解析で得られるダイナミクスの幾何的特徴をある程度おさえることが可能である。よって、今後の議論では、このD関数も系の解析に併用することにする。

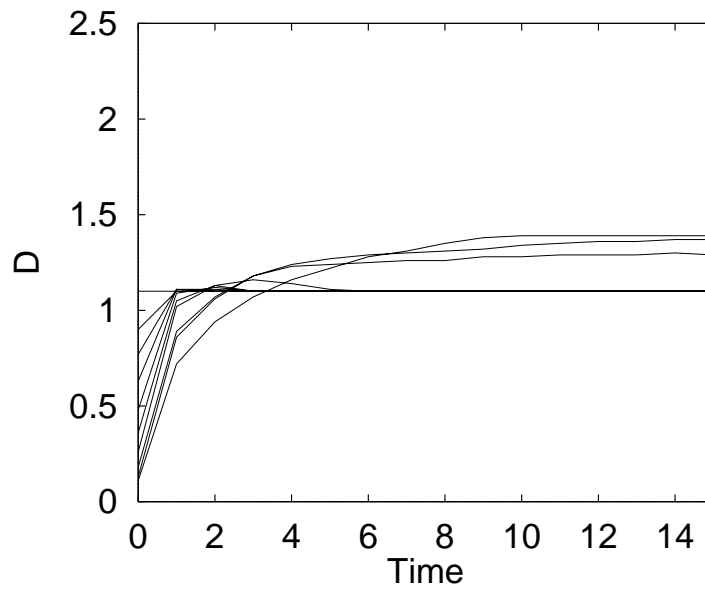


図 3.22: 自己相関連想記憶の想起過程における D 関数の変化 (自己結合なし, $a = 0.1, N = 500$)。

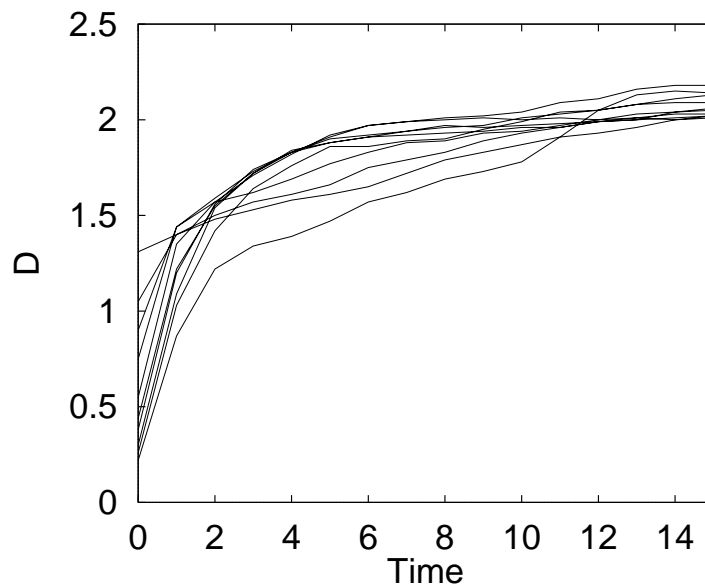


図 3.23: 自己相関連想記憶の想起過程における D 関数の変化 (自己結合なし, $a = 0.2, N = 500$)。

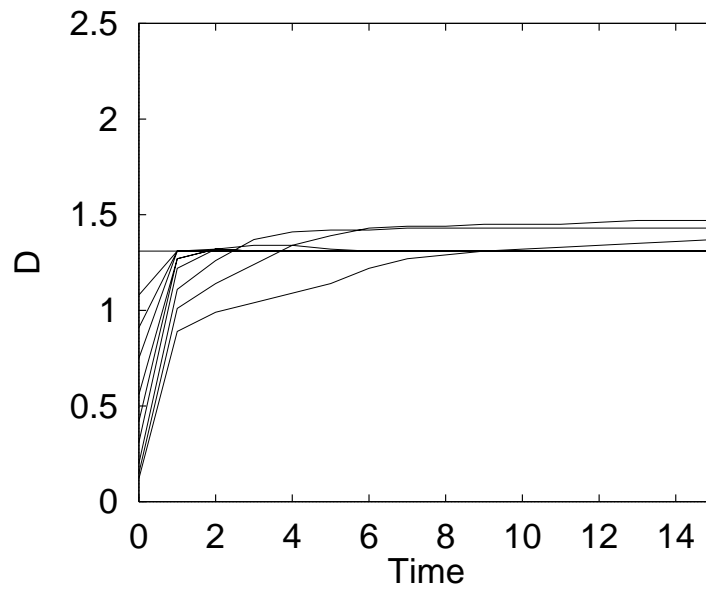


図 3.24: 自己相関連想記憶の想起過程における D 関数の変化 (自己結合あり, $a = 0.1, N = 500$)。

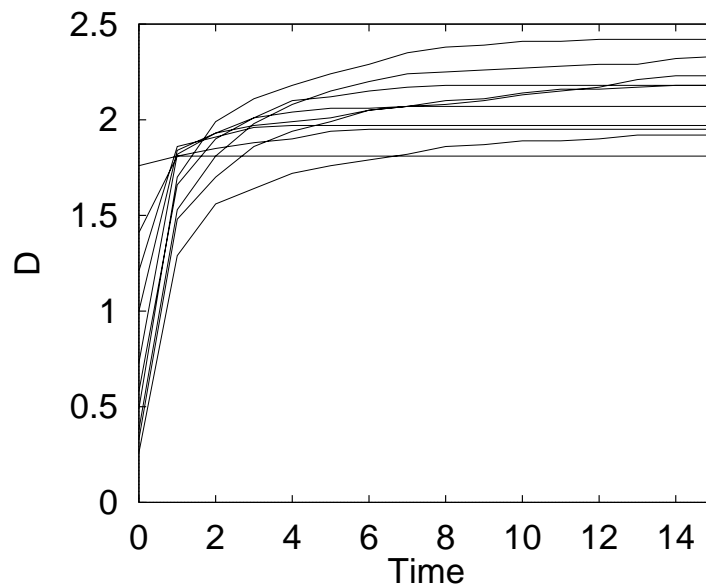


図 3.25: 自己相関連想記憶の想起過程における D 関数の変化 (自己結合あり, $a = 0.2, N = 500$)。

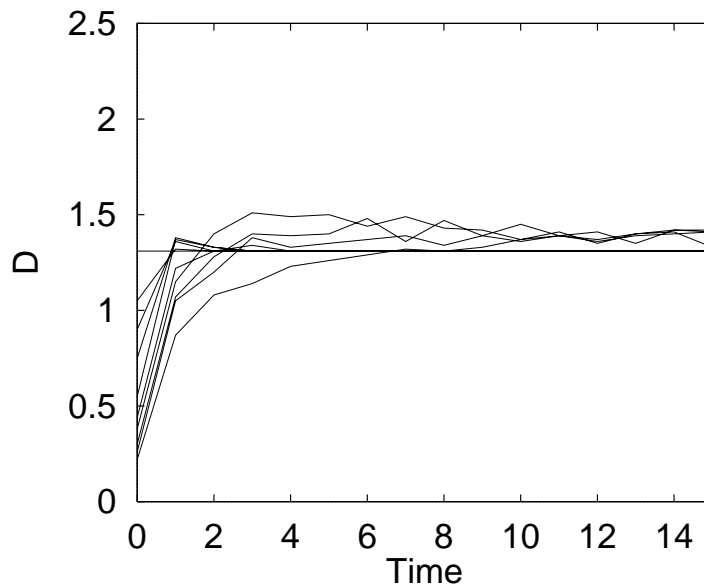


図 3.26: 自己相関連想記憶の部分反転法による想起過程での D 関数の変化 (想起可能時, $a = 0.2, N = 500, h = 1.5, \lambda = 1.0, \kappa = 1$)。

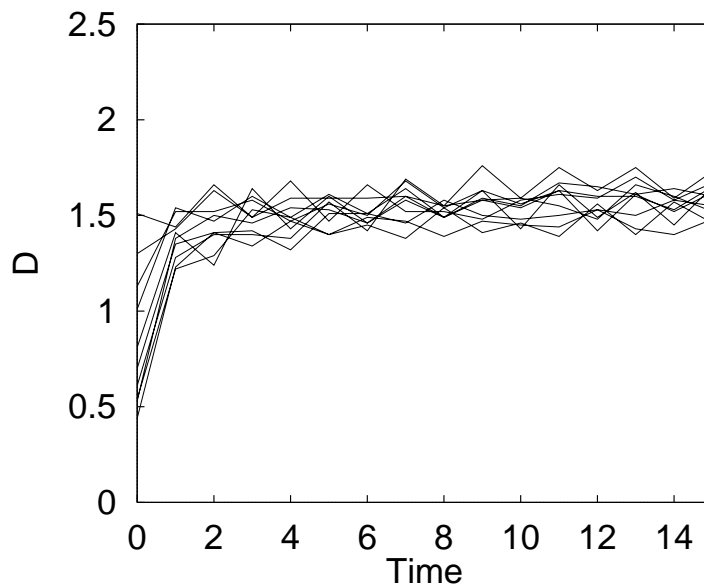


図 3.27: 自己相関連想記憶の部分反転法による想起過程での D 関数の変化 (想起不能時, $a = 0.4, N = 500, h = 1.5, \lambda = 1.0, \kappa = 1$)。

3.2 記憶空間の分割

自己相関連想記憶においては、固有値の分布の広がり記憶を不安定にさせる。自己結合を導入すると系を安定にできるが、今度は引き込み領域が失われる。このトレードオフを回避し、自己相関記憶行列の記憶容量を向上させるためには、自己結合の導入なしに固有値の分布を狭く保つ学習法を考える必要がある。これを実現する方法として、この節では符号反転記憶法とペアリング記憶法を導入する [2, 4]。

3.2.1 符号反転記憶法の導入

符号反転記憶法は、記憶パターンの自己相関行列にかかる符号を交互に変えて記憶する方法である。このとき、記憶行列は

$$w_{ij}^{\pm} = \frac{1}{N} \sum_{\mu=1}^P (-1)^{\mu} s_i^{\mu} s_j^{\mu} \quad (3.24)$$

で表される。これを符号反転記憶行列と呼ぶ。つまり、従来のモデルでは各記憶パターンの自己相関行列を全て足し合わせていたところを、符号反転記憶法ではその半分について -1 を掛けて足すことにする。これにより、記憶パターンどうしで自己結合を打ち消し合うようになり、自己結合は $-1/N$ または 0 に制限され、事実上自己結合は無視できるようになる。

符号反転記憶法において -1 を掛けて記憶したパターンは、記憶パターンとその反転パターンの間を振動することになる。従来の方法で記憶した場合には、パターン \mathbf{s} を記憶すると、その反転パターン $-\mathbf{s}$ も付随して記憶することになることを思い起こせば、この二つのパターン間を振動する状態であっても、記憶パターンに「収束」したと考えてもよいであろう。振動状態を嫌うのであれば、想起のダイナミクスを

$$x_i(t+1) = \operatorname{sgn}\left(\sum_{j=1}^N w_{ij} \operatorname{sgn}\left(\sum_{k=1}^N w_{jk} x_k(t)\right)\right) \quad (3.25)$$

とすればよい。つまり、従来の想起のダイナミクスの 2 ステップを 1 ステップであると考えることにすれば、観測される想起状態は振動的でなくなる。

3.2.2 符号反転記憶法の利点

以下の議論では、まず、符号反転記憶法による固有値の分布を調べ、符号反転記憶行列の特徴を明らかにする。次に、記憶パターンベクトルの固有空間解析を行う。最後に、符号反転記憶法における記憶パターンの安定性と引き込み領域についての実験結果を示す。

上で従来のモデルについて行った実験と同様に $P = 60$ 個のランダムパターンを記憶パターンとして、その符号反転記憶行列の固有値の分布を調べた結果を図 3.28 に示す。

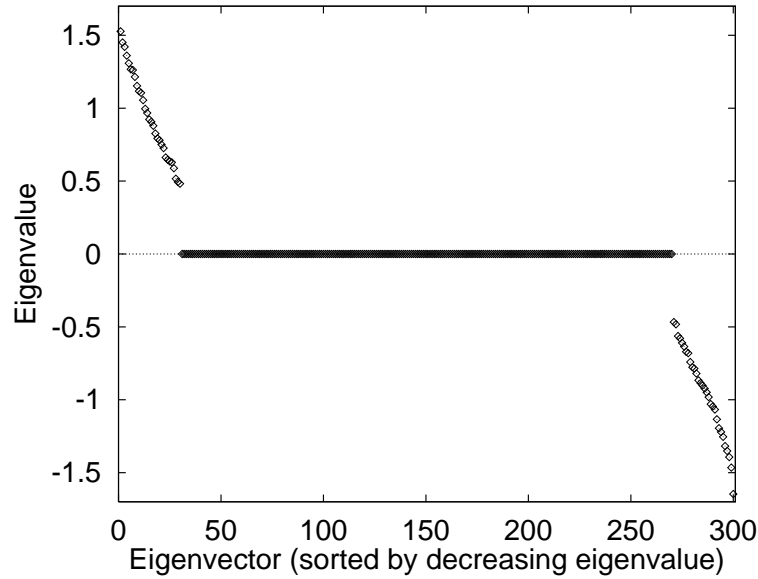


図 3.28: 符号反転記憶行列 W^\pm の固有値 ($N = 300$)。

この図を図 3.4 と比較すると、符号反転記憶法は固有値の絶対値の分布の広がりを抑え、 $P = 60$ でも通常の記憶法の $P = 30$ 程度の分布幅を実現していることが分かる。

次に、記憶パターンを固有ベクトルで展開した結果を図 3.29 に示す。この図から、記憶パターンは同符号の固有値をもつ固有ベクトルから構成されることが読みとれる。すなわち、空間が正の係数で記憶したパターンと負の係数で記憶したパターンとで分離され、記憶の二重化が実現されている。

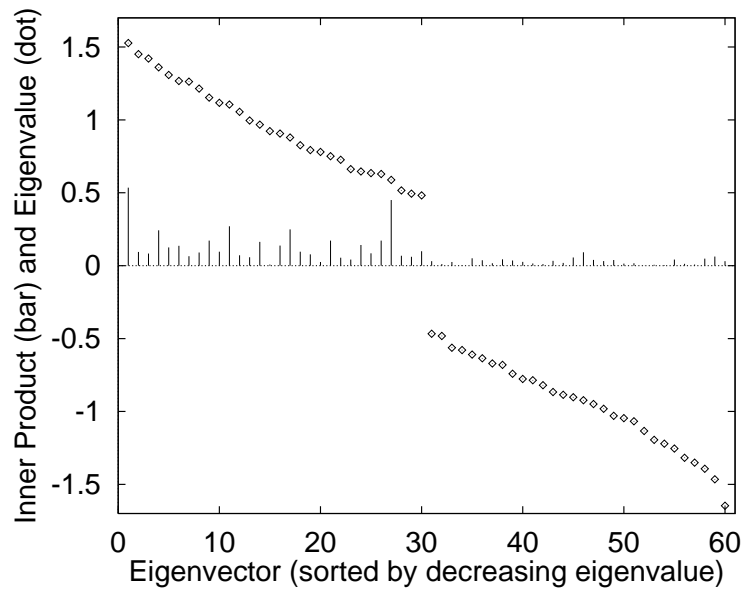
この記憶二重化のメカニズムは次のように理解される。簡単のため、ここでは、二つの記憶パターンを、従来の記憶法、及び符号反転記憶法で記憶する例をとりあげて説明する。今、記憶する二つのパターンを \mathbf{s}^1 , \mathbf{s}^2 とする。また、 $W^1 = \mathbf{s}^1 \mathbf{s}^{1T}$, $W^2 = \mathbf{s}^2 \mathbf{s}^{2T}$ とする。従来の記憶法では、記憶行列は $W^s = W^1 + W^2$ 、符号反転記憶法では $W^\pm = -W^1 + W^2$ となる。それぞれの行列の固有ベクトルを計算すると、 W^s の固有ベクトル $\mathbf{e}^{s(1)}$, $\mathbf{e}^{s(2)}$ は

$$\begin{aligned} \mathbf{e}^{s(1)} &= \mathbf{s}^1 + \mathbf{s}^2 \\ \mathbf{e}^{s(2)} &= \mathbf{s}^1 - \mathbf{s}^2 \end{aligned} \quad (3.26)$$

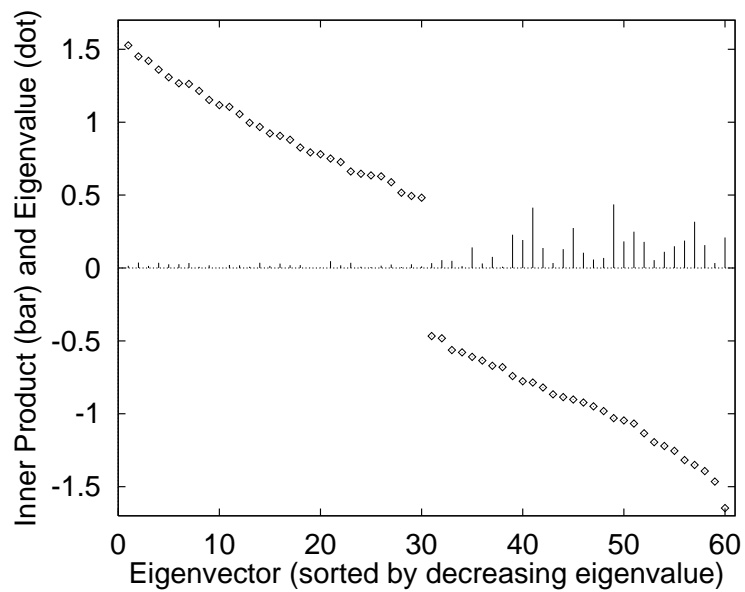
となり、 W^\pm の固有ベクトル $\mathbf{e}^{\pm(1)}$, $\mathbf{e}^{\pm(2)}$ は

$$\mathbf{e}_j^{\pm(i)} = \begin{cases} |\mathbf{s}_1 - \mathbf{s}_2|^{\frac{1}{2}} s_j^i, & \text{if } s_j^1 = s_j^2 \\ |\mathbf{s}_1 + \mathbf{s}_2|^{\frac{1}{2}} s_j^i, & \text{if } s_j^1 \neq s_j^2 \end{cases} \quad (3.27)$$

となる。従来の記憶法では記憶パターンを固有ベクトル展開するとその係数の絶対値は等しく、記憶パターンは二つの固有ベクトルで分担して表現される。特に、 \mathbf{s}_1 と \mathbf{s}_2 が直交する場合は、固有値が等しくなるので、記憶パターンを張る空間上ではパターンは全



(a) 正の記憶パターンと W^\pm の固有ベクトルの内積と固有値



(b) 負の記憶パターンと W^\pm の固有ベクトルの内積と固有値

図 3.29: 記憶パターンを符号反転記憶行列 W^\pm の固有ベクトルで展開した結果 ($a = 0.2, N = 300$)。

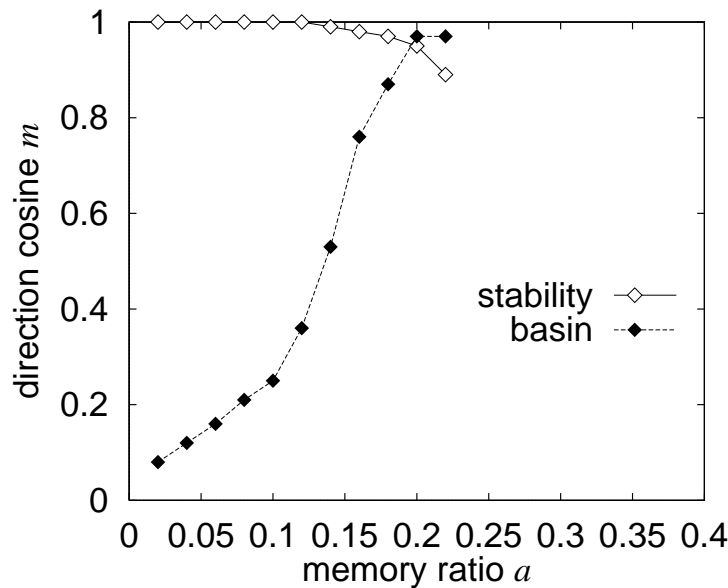


図 3.30: 符号反転記憶法を用いたときの記憶パターンの安定性と引き込み領域 ($N = 500$)。

く動かない。一方、符号反転記憶法においては記憶パターンをできるだけ記憶時に自己相関行列に掛ける符号と同一符号の固有値をもつ固有ベクトルに対応させるように表現される。特に、 s_1 と s_2 が直交する場合は

$$|s_1 - s_2| = |s_1 + s_2|$$

であるので、記憶パターンが固有ベクトルに完全に対応する。このように、符号反転記憶法では、パターン空間が正で記憶されたパターンの空間と負で記憶されたパターンの空間に分離される。よって、多数のパターンを記憶する場合、この空間の分離によって記憶が多重化される。

このことから、符号反転記憶法が従来の記憶法より記憶容量を上昇させることが予想される。実際、数値実験により符号反転記憶法と従来の記憶法の記憶パターンの安定性と引き込み領域の大きさを比較した結果を図3.30に示す。

まず、自己結合のあるモデルと符号反転記憶法を比較してみる。引き込み領域について比較すると、符号反転記憶法の方が遥かに広がっている。これは、上で述べたように、自己結合のあるモデルでは、線形変換が全体として淀んでいるためであると説明できる。続いて、記憶したパターンの安定性を比較する。記憶パターンの安定性は線形変換による変化量で決まるから、固有値の最大値と最小値の比を考えれば、定性的な評価をすることができる。ここで、既に挙げた固有値の分布の数値実験の結果における記憶空間を張る固有ベクトルの最大固有値と最小固有値の比を比べると、自己結合があるモ

デルで記憶率が0.1の場合と、符号反転記憶法で記憶率が0.2の場合がほぼ等しくなる。このことから、符号反転記憶法は、記憶したパターンの安定性という点でも2倍の記憶容量を実現することが想像される。にもかかわらず、図3.30を比較すると明らかなように、自己結合のあるモデルの方が記憶パターンの固定が強くなっている。この理由としては、符号反転記憶法において、空間が正で記憶したものと負で記憶したものに完全には分離されていない(図3.28)ことが、記憶パターンを不安定化する要因となっているためと考えられる。

つぎに、自己結合のないモデルと符号反転法を比較すると、引き込み領域の広さについては符号反転記憶法が大きく上回る。この引き込み領域の広さは、パターンを張らない空間の固有値を0に抑えていることに起因すると考えられる。また、記憶パターンの安定性は、符号反転記憶法によってかなり改善されている。上で述べたように、自己結合のあるモデルは、その記憶空間を張る固有空間中の最大固有値/最小固有値に対する比が大きいものの、符号反転記憶法において記憶空間が完全に分離していないため、符号反転法より記憶パターンが安定性していた。しかし、自己結合のないモデルでは、全ての固有値から a を引いたため、最大固有値/最小固有値の比がさらに大きくなり、その結果、符号反転記憶法より記憶パターンが不安定化するに至ると考えられる。

3.2.3 ペアリング記憶法

次に、符号反転記憶法と同じ能力を持つペアリング記憶法について説明する。ペアリング記憶法とは記憶行列を

$$w_{ij}^p = \frac{1}{N} \sum_{\nu=1}^{\frac{P}{2}} (s_i^{2\nu-1} s_j^{2\nu} + s_i^{2\nu} s_j^{2\nu-1}) \quad (3.28)$$

のように構成する方法である。

この記憶法で得られた記憶行列を使うと、ペアにしたパターンが交互に想起される。つまり、周期2の時系列を相互相関行列で記憶することに対応する。しかし、想起ダイナミクスとして式(3.25)を考えれば、この系も自己連想記憶システムとして機能する。この式(3.25)で与えられる想起ダイナミクスのもとでは、ペアリング記憶法は自己連想記憶として符号反転記憶法と同等の性質を持つ。以下に、ペアリング記憶法が作り出す幾何構造を調べることにする。

ペアリング記憶行列の固有空間を調べた結果を図3.31と図3.32に示す。ペアリング記憶では、固有値分布が符号反転記憶法と同じ形になり、記憶パターン空間の固有値の絶対値の広がり通常は自己相関連想記憶の半分に抑えられている。しかし、記憶パターンは正の固有値をもつ空間と負の固有値をもつ空間の両方にほぼ均等に成分を持っており、符号反転記憶のような空間の分割は行われていない。

ペアリング記憶法で二つのパターンを記憶する場合を例にとり、この記憶行列の固有

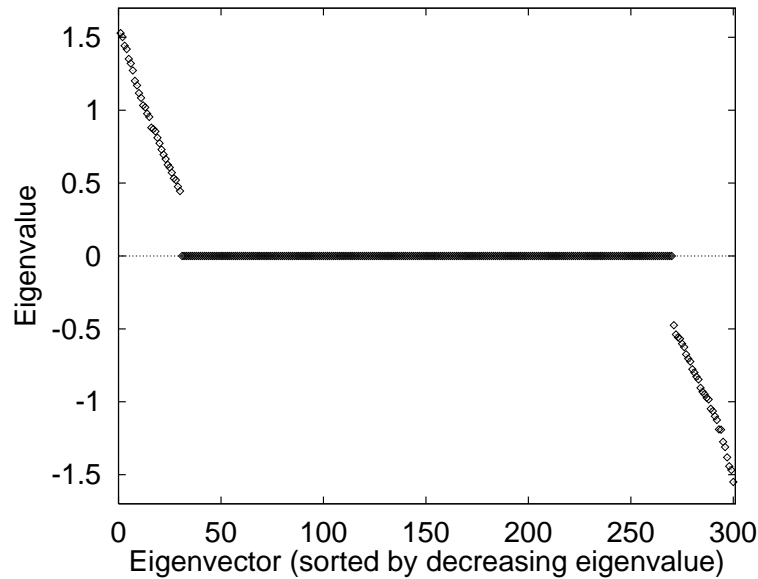


図 3.31: ペアリング記憶行列 W^p の固有値 ($N = 300$)。

空間の構成を詳しく見る。この行列の固有値ベクトル $e^{p(1)}, e^{p(2)}$ は、

$$\begin{aligned} e^{p(1)} &= s^1 + s^2 \\ e^{p(2)} &= s^1 - s^2 \end{aligned} \quad (3.29)$$

で、これは W^s の固有ベクトルと全く同じである。しかしながら、その固有値 λ_1, λ_2 はそれぞれ正負に分かれる。よって、 s^1, s^2 の共通成分は固定し、非共通成分は振動する形になっている。ここで、注目すべきことは、このペアリング記憶行列は、 $(s^1 + s^2)$ と $(s^1 - s^2)$ を符号反転記憶法で記憶した行列と同じである。このことから、符号反転記憶法と同様に、記憶パターン空間の固有値の絶対値の分散を小さく抑えることができていると考えられる。

符号反転記憶法を用いた場合とペアリング記憶法を用いた場合について、記憶したパターンの想起ダイナミクスにおける安定性と、引き込み領域の広さを数値実験で比較した結果を図 3.33 に示す。このように、ペアリング記憶法は符号反転記憶法と同等の連想能力を達成していることが分かる。

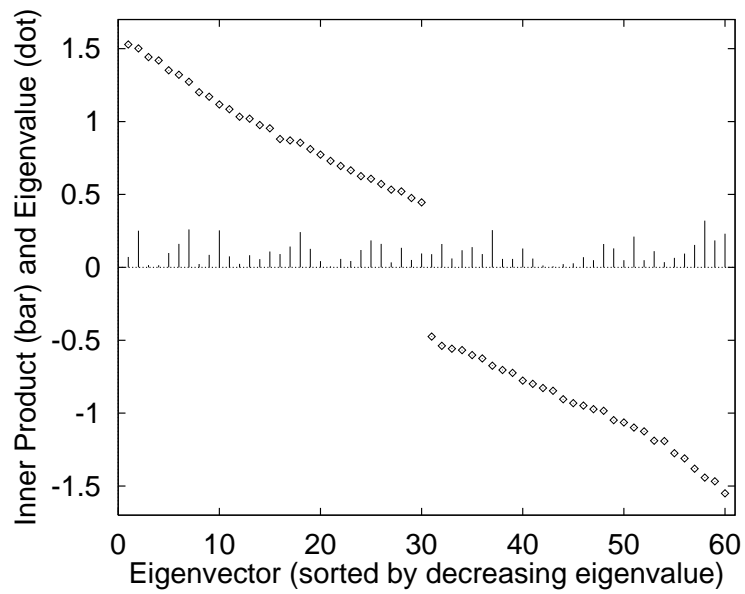


図 3.32: 記憶パターンをペアリング記憶行列 W^p の固有ベクトルで展開した結果 ($a = 0.2, N = 300$)。

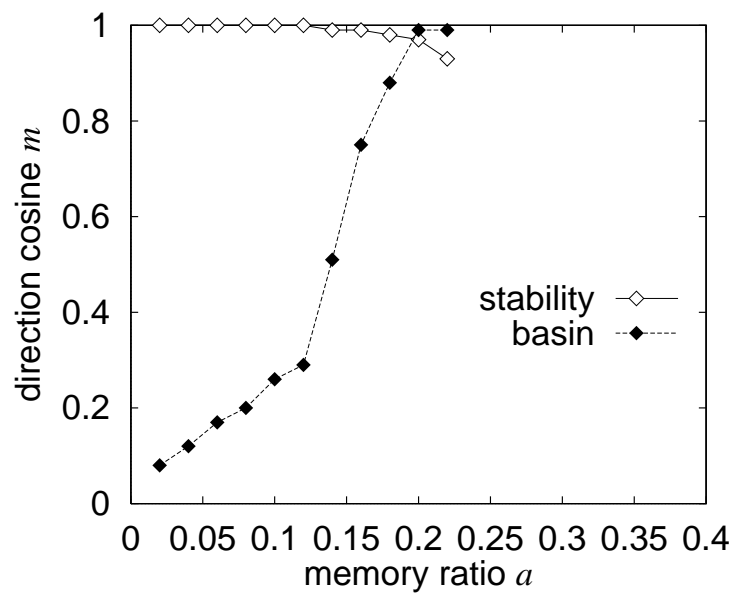


図 3.33: ペアリング記憶法を用いたときの記憶パターンの安定性と引き込み領域 ($N = 500$)。

3.2.4 記憶の分割とヒステリシス

既に指摘したように、正の記憶空間と負の記憶空間が完全に分離されていないことが、符号反転記憶法における記憶パターンの不安定性の要因であった。この小節では、想起ダイナミクスにヒステリシスを入れることにより、記憶パターンを安定化することを考える。

ヒステリシスは数理的に自己結合と完全に等価である。ヒステリシスを入れたニューロンを用いた連想記憶の性質については、既に理論的な研究が行われている [42]。ヒステリシスニューロンでは、状態遷移が鈍り、すべてのパターンが安定化へと向かうので、記憶パターンは安定化されるが、逆に引き込み領域は小さくなる。

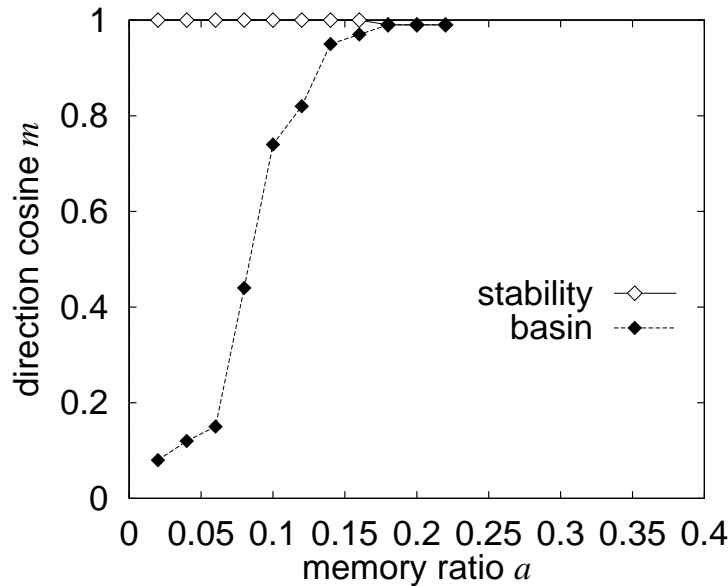


図 3.34: 符号反転記憶にヒステリシスダイナミクスを用いたときの記憶パターンの安定性と引き込み領域 ($N = 500$)。

符号反転記憶行列での想起で、ヒステリシスと同じ効果をもたせるには、

$$x_i(t+1) = \text{sgn}\left(\sum_{j=1}^N w_{ij}x_j(t) + \varepsilon a \psi(\mathbf{x}(t))x_i(t)\right) \quad (3.30)$$

という想起ダイナミクスを用いればよい [9]。ここで

$$\psi(\mathbf{x}) = \text{sgn}\left(\sum_{i=1}^N \left(\sum_{j=1}^N w_{ij}x_j\right)x_i\right) \quad (3.31)$$

である。このダイナミクスにおいては、 $\frac{\pi}{2}$ 以上回転するパターンは負の記憶パターンに近いとみなしてより振動的になるようにし、逆に $\frac{\pi}{2}$ 以下しか回転しないパターンは正の記憶パターンに近いとみなして、より固定的になるようにしている。この方法によって、記憶空間の分離が完全でないのを補う効果が得られると期待できる。このダイナミクスを使った実験の結果を図 3.34 に示す。実験の結果、予想通り、記憶パターンの安定性は増し、引き込み領域は減ることがわかる。ここで注目すべき点として、自己結合のある普通の自己相関行列と同じ安定性を符号反転記憶法をヒステリシスダイナミクスで想起する場合にもたせるには、通常モデルの自己結合に比べて 4 分の 1 から半分程度の強さのヒステリシスをいれればよいことである。また、その時、引き込み領域は符号反転記憶の方が広い。

3.2.5 記憶の分割と部分反転法

通常の自己相関連想記憶においては、部分反転法によって記憶容量を向上させることができた。

しかし、部分反転法をそのまま符号反転記憶とペアリング記憶に適用すると、連想記憶の記憶容量を向上させることができない。これは、部分反転法が振動的なパターンに対応することができないことによる。そこで、以下に、新しい記憶法による連想記憶の能力を改善するようなダイナミクスを提案する。

符号反転記憶、ペアリング記憶の両方に適用できるダイナミクスの改良として、式 (3.25) で表されるダイナミクスについて部分反転法を使用する方法が考えられる。ここで、反転の入れ方に次のような工夫が必要である。内部状態 u_i の絶対値が大きいニューロンの影響力を弱める方法として、単純な相関学習で得られた記憶行列を用いる場合は、式 (3.7) のように、記憶行列 $W = [w_{ij}]$ を掛けていた。しかし、式 (3.25) で与えられるダイナミクスにおいては、二度の W による変換が含まれているので、 $|u_i|$ の大きなニューロンの影響が反転するようにするには、 $V = W \times W$ を掛ける必要がある。よって、

$$x_i(t+1) = \text{sgn}\left(u_i(t) - \sum_{j=1}^N v_{ij} \phi(u_j(t))\right) \quad (3.32)$$

で表されるダイナミクスを使用することによって、連想記憶の記憶容量を向上できることが予想される [4]。ここで、

$$u_i(t) = \sum_{j=1}^N w_{ij} \text{sgn}\left(\sum_{k=1}^N w_{jk} \text{sgn}(x_k(t))\right) \quad (3.33)$$

で、また、

$$v_{ij} = \sum_{k=1}^N w_{ik} w_{kj} \quad (3.34)$$

である。

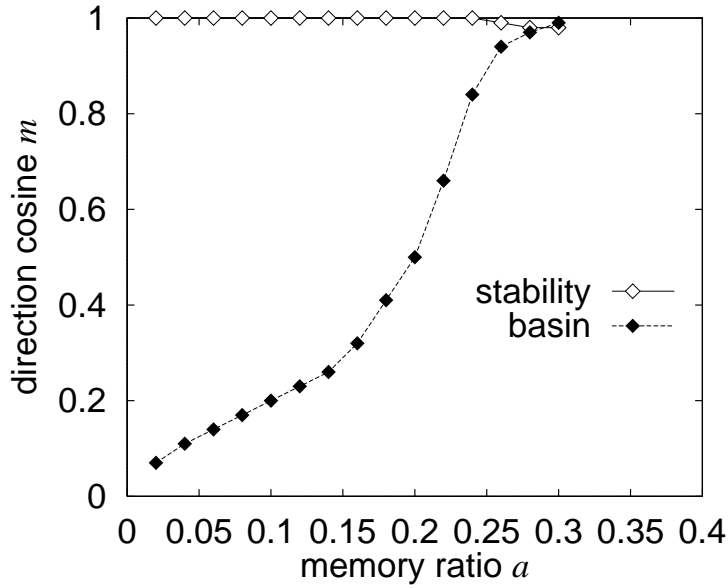


図 3.35: 符号反転記憶に部分反転法を用いたときの記憶パターンの安定性と引き込み領域 ($N = 500$)。

符号反転記憶について、通常の想起法を使用した時と、部分反転法を応用した想起法を使用した場合の記憶パターンの安定性と引き込み領域について、数値実験で得られた結果を図 3.35 に示す。この実験結果と、図 3.35 の結果を比較すると、式 (3.32) で表される部分反転ダイナミクスが、予想通り、符号反転記憶を使用した連想記憶の記憶パターンの安定性と引き込み領域を向上させることがわかる。尚、ペアリング記憶にここで提案した部分反転法の変形を応用すると同様の安定性と引き込み領域が得られることが確かめられている。

振動に対処する部分反転法として、上記で述べた方法の他に、

$$x_i(t+1) = \text{sgn}\left(u_i(t) - \sum_{j=1}^N w_{ij}\phi(\tilde{u}_j(t))\right) \quad (3.35)$$

で表されるダイナミクスを使用することによって、連想記憶の記憶容量を向上できることが予想される [21]。ここで

$$u_i(t) = \sum_{j=1}^N w_{ij}x_j(t) \quad (3.36)$$

$$\tilde{u}_i(t) = \sum_{j=1}^N v_{ij}x_j(t) \quad (3.37)$$

$$v_{ij} = \sum_{k=1}^N w_{ik}w_{kj} \quad (3.38)$$

である。この方法は、相互相関連想記憶においても有効な方法であるので、次章の非対称荷重を持つネットワークの議論において詳しく述べることにする。

3.3 まとめ

本章においては、第2章で論じたニューロダイナミクスの幾何的視点をもとに、対称な荷重行列を持つニューラルネットワークの離散時間ダイナミクスの性質について述べた。対称な荷重行列として持つネットワークの例として、自己相関連想記憶をとりあげ、その力学の幾何構造を明らかにした。

対称な荷重行列を持つニューラルネットワークでは、系は荷重行列の固有ベクトルでその固有値の絶対値が大きい方向へ状態遷移が進む。その状態遷移の速度は固有値分布の広さに依存し、固有値の分布が狭いと状態遷移の速度は小さく、非線形の安定化効果で状態遷移は途中で停止する。

自己相関連想記憶においては、記憶パターンは記憶したパターン数の次元を持つ記憶空間の固有ベクトル群全体で張られている。記憶パターン数が増え、この記憶空間の固有値の分布が広がると、記憶パターンは不安定になり、大きい固有値を持つ固有ベクトル方向へと状態遷移が進む。これが記憶容量を超えた時に生じる偽記憶のメカニズムである。部分反転法は、内部状態の大きなニューロンの出力の影響を抑えることで、大きい固有値を持つ固有ベクトルへの状態遷移を抑える。これが部分反転法が記憶パターンを安定化し、記憶容量を増大させている。

また、幾何的視点に基づき、記憶パターン数の増大による固有値の分布の広がりを抑える新たな記憶法を提示した。提案した符号反転記憶法とペアリング記憶法は、記憶空間を二分割することで固有値分布の広がりを抑え、記憶容量の向上を実現することが示された。

次章においては、荷重行列が非対称の場合に解析対象を拡張して、ニューロダイナミクスの幾何を論じることにする。

第 4 章：非対称ネットワークの幾何

前章では、対称な結合荷重をもつ系について、ニューロダイナミクスの幾何構造を調べた。この章においては、結合荷重が非対称な場合について、荷重行列の固有空間解析に基づく幾何的解析を行う。まず、第1節では、非対称行列の固有空間解析の方法について述べる。次に、第2節では、非対称荷重を持つニューラルネットワークの例として相互関連連想記憶を取り上げ、その幾何構造を解析する。最後に、第3節において、非対称のネットワーク特有の複雑な動的振舞について幾何的視点から説明を行うと同時に、非対称ネットワークに対応する部分反転法を提案する。

4.1 非対称行列の幾何構造

この節では、非対称な結合荷重をもつニューラルネットワークを解析する基盤となる非対称行列の固有空間解析の基礎について述べる。

以下の議論では、無印のベクトル \mathbf{v} は縦ベクトルとし、その複素共役を $\tilde{\mathbf{v}}$ 、転置を \mathbf{v}^T と表す。非対称行列 W には、右固有ベクトル $\mathbf{f}^\mu = (f_1^\mu f_2^\mu \cdots f_N^\mu)^T$ と左固有ベクトル $\mathbf{g}^{\mu T} = (g_1^\mu g_2^\mu \cdots g_N^\mu)$ が存在する。すなわち、

$$\lambda_\mu \mathbf{f}^\mu = W \mathbf{f}^\mu \quad (4.1)$$

$$\lambda_\mu \mathbf{g}^{\mu T} = \mathbf{g}^{\mu T} W \quad (4.2)$$

が満たされている。ここで、 λ_μ はそれぞれの固有ベクトルに対応する固有値である。行列 W は、これらの固有値と固有ベクトルを使って

$$W = \sum_{\mu=1}^N \lambda_\mu \mathbf{f}^\mu \mathbf{g}^{\mu T} \quad (4.3)$$

と表すことができる。ここで、

$$\mathbf{g}^{\mu T} \mathbf{f}^\nu = \delta_{\mu\nu} \quad (4.4)$$

の関係が成り立つ ($\delta_{\mu\nu}$ はクロネッカーのデルタ関数)。

W が非対称行列の場合は、固有値に複素数が現れる。但し、 W の成分が全て実数の場合は、共役複素数がペアになって現れる。つまり、 λ が固有値なら $\tilde{\lambda}$ も固有値である。今、 W の固有値のうち、実数固有値の集合を $\lambda_\mu \in \Lambda_R$ とする。また、複素固有値については複素共役のペア $(\lambda_\nu, \lambda_{\nu'})$ の組を $G_{\nu, \nu'}$ とし、その集合を $G_{\nu, \nu'} \in \Lambda_C$ とする。すると、式 (4.3) は

$$W = \sum_{\Lambda_R} \lambda_\mu \mathbf{f}^\mu \mathbf{g}^{\mu T} + \sum_{\Lambda_C} (\lambda_\nu \mathbf{f}^\nu \mathbf{g}^{\nu T} + \lambda_{\nu'} \mathbf{f}^{\nu'} \mathbf{g}^{\nu' T}) \quad (4.5)$$

と変形できる。ここで、複素固有値をもつ固有ベクトルについて、 $f_i^\nu = |f_i|e^{i\phi_i^\nu}$, $g_j^\nu = |g_j|e^{i\psi_j^\nu}$ とおき、 $\lambda_\nu, \lambda_{\nu'}$ がそれぞれ $r_\nu e^{i\omega_\nu}, r_{\nu'} e^{-i\omega_{\nu'}}$ と書き換えられることに着目すると、式 (4.5) の行列 W の (i, j) 成分 w_{ij} は

$$w_{ij} = \sum_{\Lambda_R} \lambda_\mu f_i^\mu g_j^\mu + \sum_{\Lambda_C} 2r_\nu |f_i^\nu| |g_j^\nu| \cos(\phi_i^\nu + \psi_j^\nu + \omega_\nu) \quad (4.6)$$

となる。更に、 W の Q 乗の成分 w_{ij}^Q は

$$w_{ij}^Q = \sum_{\Lambda_R} \lambda_\mu^Q f_i^\mu g_j^\mu + \sum_{\Lambda_C} 2r_\nu^Q |f_i^\nu| |g_j^\nu| \cos(\phi_i^\nu + \psi_j^\nu + Q\omega_\nu) \quad (4.7)$$

となる。

ここで、 W による線形変換を繰り返す、式

$$\mathbf{v}(t+1) = \frac{\sqrt{N}W\mathbf{v}(t)}{|W\mathbf{v}(t)|} \quad (4.8)$$

で表される時間発展について考える。初期状態 $\mathbf{v}(0)$ の成分が

$$v_i(0) = \sum_{\Lambda_R} C_\mu f_i^\mu + \sum_{\Lambda_C} 2C_\nu |f_i^\nu| \cos(\phi_i^\nu + \zeta^\nu) \quad (4.9)$$

で表されるとすると、 t ステップ後の状態 $\mathbf{v}(t)$ は

$$\mathbf{v}(t) = \frac{\sqrt{N}W^t\mathbf{v}(0)}{|W^t\mathbf{v}(0)|} \quad (4.10)$$

であり、その成分を右固有ベクトルで展開すると

$$v_i(t) = K_t \left(\sum_{\Lambda_R} \lambda_\mu^t C_\mu f_i^\mu + \sum_{\Lambda_C} 2r_\nu^t C_\nu |f_i^\nu| \cos(\phi_i^\nu + \zeta^\nu + t\omega_\nu) \right) \quad (4.11)$$

$$K_t = \frac{\sqrt{N}}{|W^t\mathbf{v}(0)|} \quad (4.12)$$

と表される。このことから、 $t \rightarrow \infty$ では、全ての固有値のノルム (λ_μ または r_ν) 中、最大のものが $\lambda_{\hat{\mu}} \in \Lambda_R$ ならば、系は $\mathbf{f}^{\hat{\mu}}$ に収束し ($\lambda_\mu < 0$ ならば周期 2 で振動)、 $r_{\hat{\nu}} \in \Lambda_C$ ならば、系の各成分が

$$v_i(\tau) = 2|f_i^{\hat{\nu}}| \cos(\phi_i^{\hat{\nu}} + \hat{\zeta} + \tau\omega_{\hat{\nu}}) \quad (4.13)$$

に従って振動する状態に陥ることが分かる ($\hat{\zeta}$ は初期位相によって決定される定数)。

ニューラルネットワークのダイナミクスでは、この線形変換の状態遷移のフローに非線形変換による量子化が加わる。次章以降の議論では、この章で議論した固有空間解析を使うことにより、非対称の荷重行列をもつニューラルネットワークの性質を解析することにする。

4.2 相互相関連想記憶

この節においては、前節で述べた非対称行列の固有空間解析を用いて、相互相関連想記憶の幾何構造を調べる。

4.2.1 相互相関連想記憶の定式化

相互相関連想記憶では、異なる二つのパターン $\mathbf{s}^a, \mathbf{s}^b$ の相関行列 $\mathbf{s}^b \mathbf{s}^{aT}$ を荷重行列に加えていく。覚えるパターンの組の集合を $(a_\mu \rightarrow b_\mu) \in M$ とすると、荷重行列は

$$W = \sum_M \mathbf{s}^{b_\mu} \mathbf{s}^{a_\mu T} \quad (4.14)$$

と表される。これにより、 \mathbf{s}^a の入力に対して \mathbf{s}^b が想起されるようになる。このように荷重行列を構成したネットワークに

$$x_i(t+1) = \text{sgn}\left(\sum_{j=1}^N w_{ij} x_j(t)\right) \quad (4.15)$$

$$\text{sgn}(u) = \begin{cases} 1, & \text{if } u \geq 0 \\ -1, & \text{if } u < 0 \end{cases} \quad (4.16)$$

で与えられる通常のニューロダイナミクスを適用すると、記憶したパターン系列の連想が実現される。

但し、リカレントなネットワークでは、時刻 t の出力が時刻 $(t+1)$ の入力となる。よって、時刻が進んでも記憶した状態が保持されるためには、図 4.1 のように周期的なパターンの列か、連想の終点を自己相関等で安定化して締め括るパターンの列を想定する必要がある。

4.2.2 相互相関連想記憶の幾何

以下の議論では、パターンの系列を記憶させた相互相関連想記憶の荷重行列がどのような固有空間を形成するかを調べ、それに基づいて想起ダイナミクスの説明を試みる。ここでは、相互相関連想記憶の例として周期的なパターンの系列を記憶した場合を取り上げる。

まず、周期 Q が同一の系列を記憶した相互相関連想記憶について、荷重行列の固有値を調べた結果を図 4.2 に示す。この場合、固有値の位相が $\omega = 0$ と $\omega = \frac{2}{3}\pi$ 付近に集中する。つまり、周期 3 でもとに戻る位相になっている。

この荷重行列を使った想起過程における右固有ベクトルと状態ベクトルの内積の変化を図 4.3 に示す。この図のように、想起過程では記憶パターンの系列と直交する空間(ノイズ空間)は、固有値の絶対値が小さく、その成分は減少する。一方、記憶パターンの

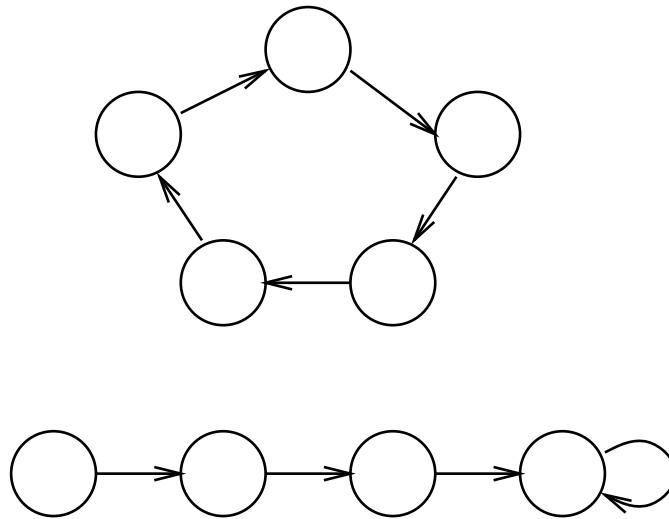


図 4.1: 相互相関連想記憶で覚える系列の種類。

系列を張る空間(記憶空間)の固有ベクトル成分は、その固有値の位相に応じた周期で振動していることが分かる。

次に、記憶した系列の周期 Q が異なる場合 ($Q = 4, 5, 6$) について荷重行列の固有値を調べた結果を図 4.4 に示す。この場合は、固有値の位相が $\omega = 0, \omega = \pi, \omega = \frac{4}{5}\pi, \omega = \frac{2}{3}\pi, \omega = \frac{1}{2}\pi, \omega = \frac{2}{5}\pi, \omega = \frac{1}{3}\pi$ 付近に出現している。これらのうち、位相が $\omega = 0, \omega = \pi, \omega = \frac{1}{2}\pi$ の回転は 4 ステップ後にもとに戻り、 $\omega = 0, \omega = \frac{4}{5}\pi, \omega = \frac{2}{5}\pi$ の回転は 5 ステップ後にもとに戻り、 $\omega = 0, \omega = \frac{2}{3}\pi, \omega = \frac{1}{3}\pi$ の回転は 6 ステップ後にもとに戻る。つまり、どれも覚えた系列の周期 4, 5, 6 のどれかでもとの位相に戻る角速度を生むようになっている。

この荷重行列を使った想起過程における固有ベクトルと状態ベクトルの内積の変化を図 4.5 に示す。ここでは、周期 4 の系列を想起する場合を示している。この図から、周期 4 の系列の想起過程では、記憶パターンの系列を張る固有ベクトルのうち、周期 4 を産み出す(固有値が $\omega = 0, \omega = \pi, \omega = \frac{1}{2}\pi$ 付近の位相をもつ)固有ベクトル成分は増加し、その固有値の位相に応じた周期で振動し、その他の固有ベクトル成分は減少していることが読み取れる。つまり、周期 Q をもつパターンの系列は、記憶空間の固有ベクトルの中で位相が $\frac{2n\pi}{Q}$ (n : 自然数) の固有値をもつ固有ベクトル群で張られる形になっている。

更に、図 4.2 と図 4.4 を比較して注目すべき点の一つがある。それは、どちらのケースも覚えたパターン数は同じであるにも関わらず、固有値の絶対値の分布は図 4.4 の方が狭

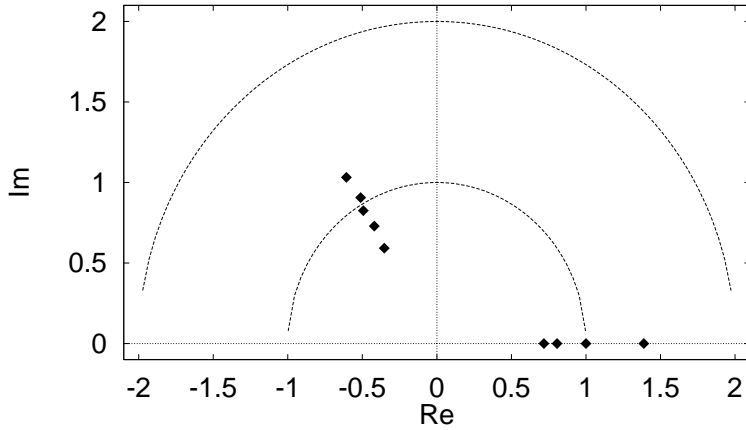


図 4.2: ニューロン数 100 のネットワークに周期 3 の系列を 5 つ覚えさせた時の固有値 (記憶空間の固有値のみを表示)。横軸と縦軸は各固有値の実部と虚部に対応している。

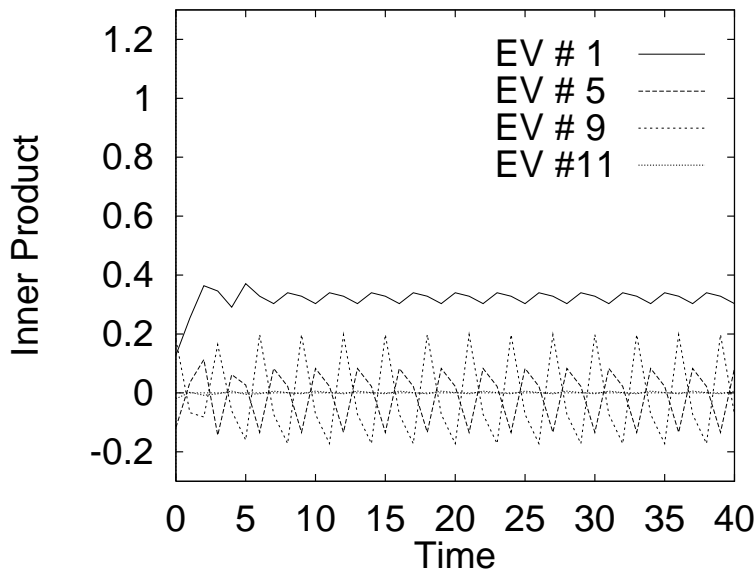


図 4.3: ニューロン数 100 のネットワークに周期 3 の系列を 5 つ覚えさせた時の想起ダイナミクス。ダイナクスにおける固有ベクトルと状態ベクトルの内積の変化をプロットしている。図中の EV# N はその絶対値が N 番目に大きい固有値をもつ固有ベクトルとの内積を意味する。ここで、固有ベクトル #11 以外は記憶空間内の固有ベクトルである。

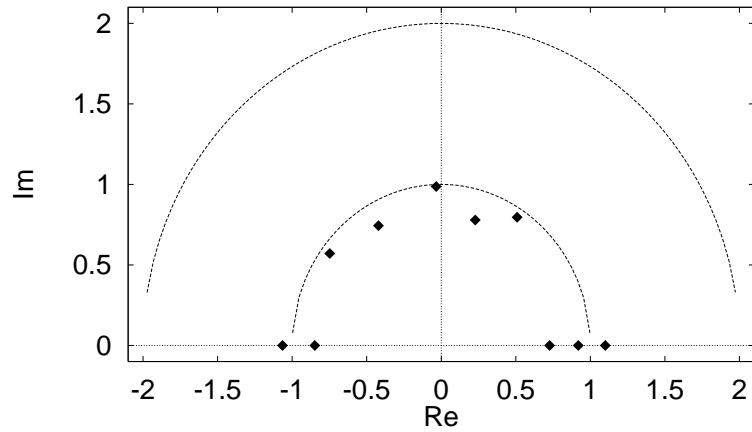


図 4.4: ニューロン数 100 のネットワークに周期 4, 5, 6 の系列を 1 つずつ覚えさせた時の固有値 (記憶空間の固有値のみを表示)。

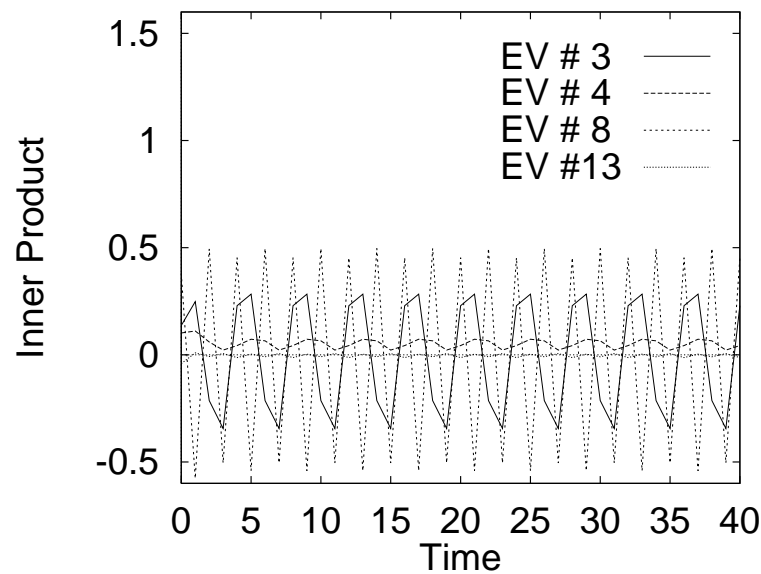


図 4.5: ニューロン数 100 のネットワークに周期 4, 5, 6 の系列を 1 つずつ覚えさせた時の想起ダイナミクスにおける固有ベクトルと状態ベクトルの内積の変化。ここでは周期 4 の系列を想起している。固有ベクトル #3 は周期 4、#4 は周期 6、#8 は周期 2 に対応する位相をもつ。固有ベクトル #13 は記憶空間外の固有ベクトルである。

くなっていることである。荷重が対称行列の場合、状態ベクトルを張る空間の固有値の絶対値の分布が小さいと、線形変換による移動量が小さくなり、状態はより安定となる。同様に、非対称結合荷重をもつネットワークでも、軌道を張る空間の固有値の絶対値の分布が狭ければ、その軌道は安定になる。周期の短い系列を記憶した場合より、周期の長い系列を記憶した場合の方が記憶容量が大きいことは既に知られた事実であるが [24]、ここで得られた固有値分布の違いはその現象に幾何的説明を与えている。

4.2.3 相互相関連想記憶の偽記憶

記憶した系列のパターン数が記憶容量を超えると、記憶系列は不安定になる。その時、系がどのような軌道に陥るかを以下に解析する。荷重行列が対称行列の場合は、大きい固有値を持つ固有ベクトル方向が安定化され、偽記憶パターンが生成される。ここでは、非対称結合のネットワークの場合も同様のことが言えるかどうかを検討する。

まず、図 4.6 に同周期の系列を記憶容量以上覚えさせた場合の固有値の分布を示す。この図から、記憶パターン数の増大が固有値分布を広げていることが分かる。また、その時現れる軌道を、固有ベクトルとの内積の変化としてプロットしたものを図 4.7 に示す。このように、対称結合荷重をもつネットワークと同様、記憶が失われる場合は絶対値の大きい固有値を持つ固有ベクトル方向へと状態ベクトルが引き寄せられていることが分かる。

また、図 4.8 と図 4.9 に周期の異なる系列を記憶容量以上に覚えさせた場合の固有値の分布とダイナミクスの様子を示す。ここでも、絶対値の大きい固有値を持つ固有ベクトルが系を支配していることが分かる。

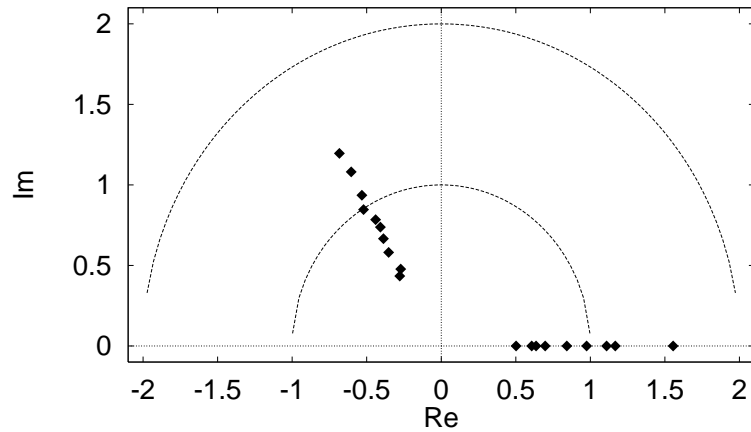


図 4.6: ニューロン数 100 のネットワークに周期 3 の系列を 10 個覚えさせた時の固有値 (記憶空間の固有値のみを表示)。

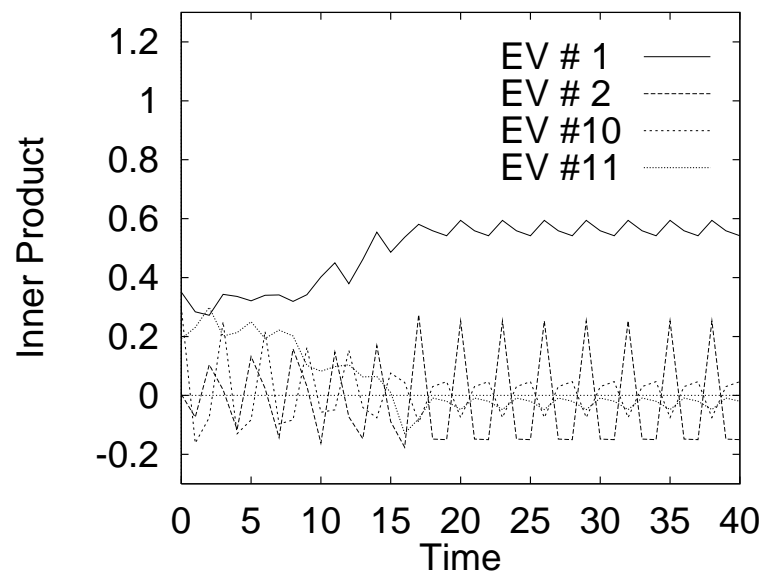


図 4.7: ニューロン数 100 のネットワークに周期 3 の系列を 10 個覚えさせた時の想起ダイナミクスにおける固有ベクトルと状態ベクトルの内積の変化。

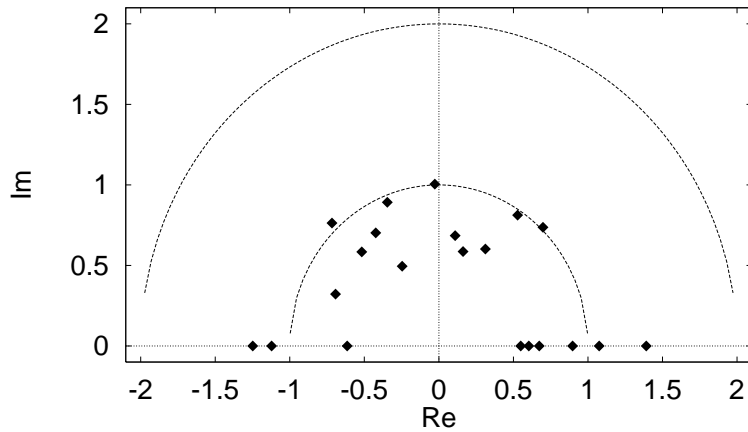


図 4.8: ニューロン数 100 のネットワークに周期 3, 4, 5, 6, 7, 8 の系列を 1 つずつ覚えさせた時の固有値 (記憶空間の固有値のみを表示)。

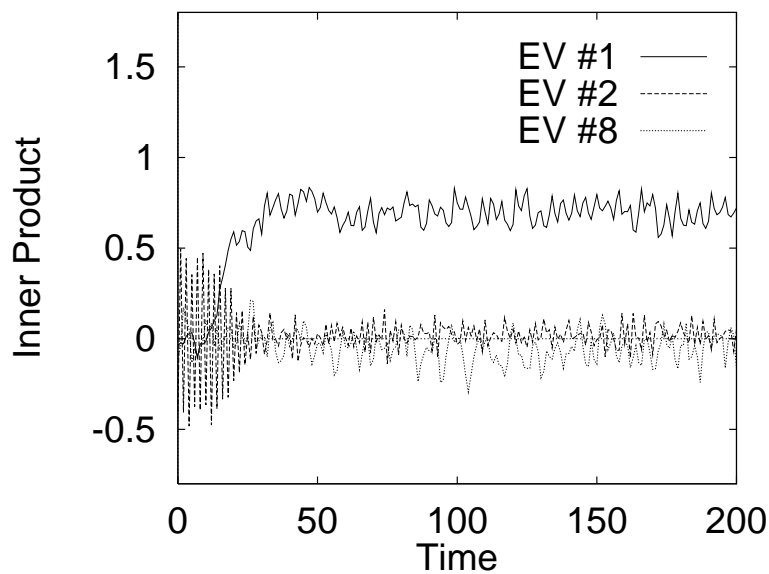


図 4.9: ニューロン数 100 のネットワークに周期 3, 4, 5, 6, 7, 8 の系列を 1 つずつ覚えさせた時の想起ダイナミクスにおける固有ベクトルと状態ベクトルの内積の変化。

4.3 複雑な動的振舞の解析

4.3.1 長周期軌道の解析

連続時間モデルの場合、非対称結合を持つネットワークの中にはカオティックな振舞を示すものがあることが知られている。離散時間モデルの場合は、状態数が有限なので、何らかの周期軌道に陥ることになるが、その周期が非常に長い軌道が生じてくる。ここでは、相互相関連想記憶とランダムネットワークで生じる長周期軌道について調べる。

相互相関連想記憶については既に前章で述べたが、連想記憶の偽記憶においても非常に長い周期軌道が生まれることがある。不安定に振動する偽記憶軌道が生じた時の固有値の分布と軌道の様子をそれぞれ図4.10と図4.11に示す。図4.11から分かるように、この場合は最大固有値が支配する期間とその次に大きい固有値が支配する期間とが交互に生じている。また、図4.10の固有値の分布から分かるように、ここでは最大固有値とその次に大きい固有値の絶対値の差は微小である。長周期軌道中に起っているのは、その両固有値をもつ固有ベクトルによる「綱引き」状態である。

次に、ランダムネットワークの固有値の分布と軌道の様子をそれぞれ図4.12と図4.13に示す。この場合は、図4.12のように固有値の位相もランダムに生じる。また、固有値の絶対値もランダムに生じ、最大固有値に近い固有値が多く存在する。その結果、それらの固有値をもつ固有ベクトルで張られる空間中を状態ベクトルがさまようことになる(図4.13)。

以上のように、非常に長い周期を持つ軌道中では、固有値の絶対値が大きい固有ベクトル中で、その値が近いものどうしが状態ベクトルを引きあっていることが分かった。

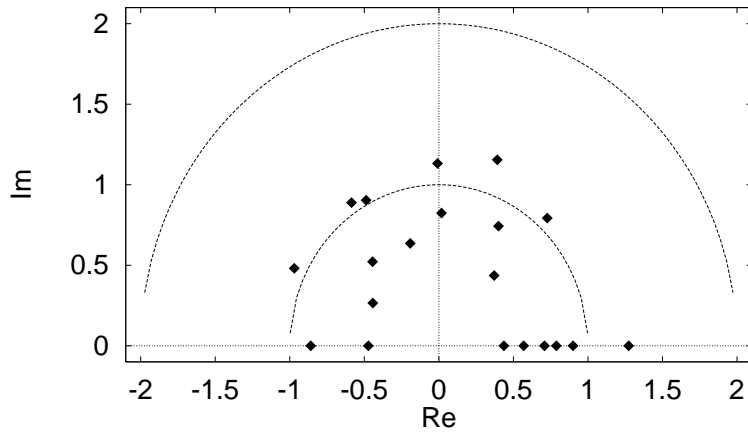


図 4.10: 相互相関連想記憶の偽記憶として長い周期軌道が現れる場合の固有値の分布(記憶空間の固有値のみを表示)。ここではニューロン数 100 のネットワーク周期 3, 4, 5, 6, 7, 8 の系列を 1 つずつ覚えさせた時の例を挙げている。

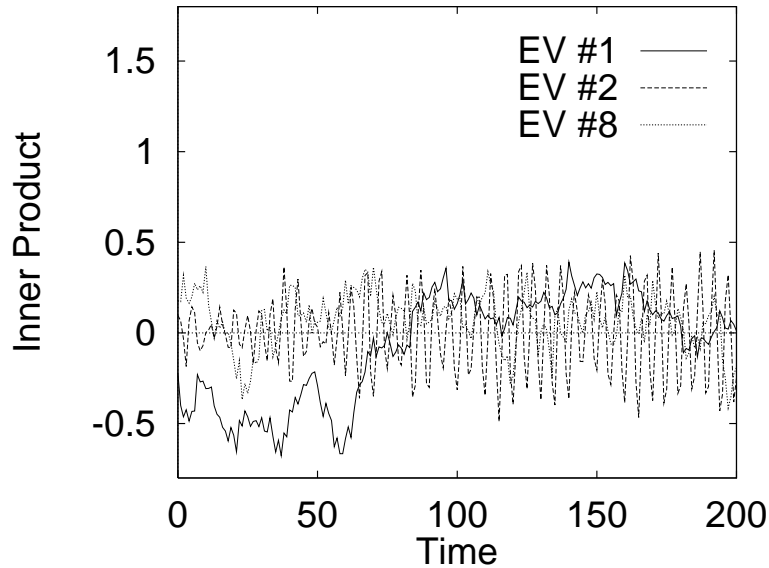


図 4.11: ニューロン数 100 のネットワークに周期 3, 4, 5, 6, 7, 8 の系列を 1 つずつ覚えさせた相互相関連想記憶において、偽記憶として生じる長い周期軌道中の固有ベクトルと状態ベクトルの内積の変化。

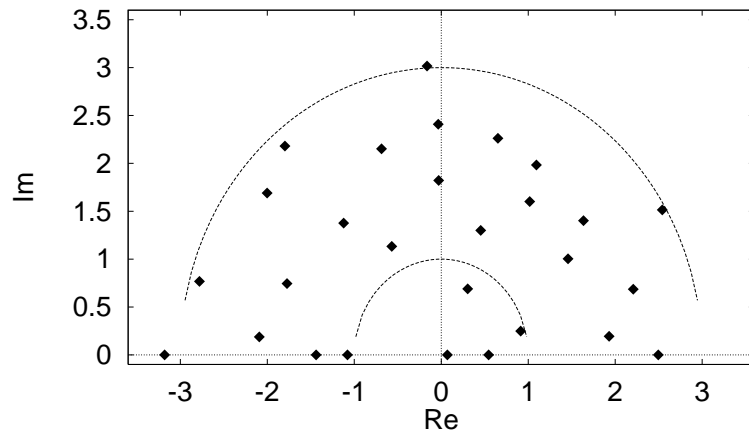


図 4.12: ランダムネットワークの固有値。ここでは各成分が $[-1,1]$ の一様分布に従う 50 次元ネットワークの全固有値を示している。

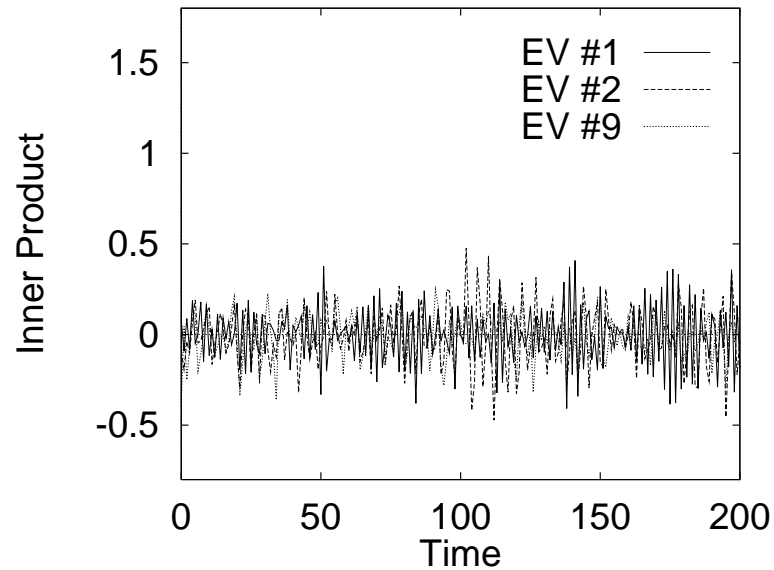


図 4.13: 各成分が $[-1,1]$ の一様分布に従う 50 次元ランダムネットワークのダイナミクスにおける固有ベクトルと状態ベクトルの内積の変化。

4.3.2 非単調ダイナミクス

対称の荷重結合をもつネットワークでは、部分反転法 [29] によって大きい固有値で張られる固有空間への流れを塞ぎ止めることができる。この方法を応用することで、記憶容量の拡大やニューロウインドウ法による選択的想起などが可能となる。しかしながら、非対称結合をもつネットワークで全く同じ方法を用いても同様の効果は得られない。ここでは、非対称結合をもつネットワークで部分反転法やニューロウインドウ法と同じ効果を得る方法について検討する。

ネットワークが対称の荷重結合を持つ場合、部分反転法のダイナミクスは、

$$x_i(t+1) = \text{sgn} \left[\sum_j w_{ij} (x_j(t) - \lambda \phi(u_j(t))) \right] \quad (4.17)$$

$$u_j(t) = \sum_{k=1}^N w_{jk} x_k(t) \quad (4.18)$$

$$\phi(u) = \begin{cases} -1, & \text{if } u < -h \\ 0, & \text{if } -h \leq u < h \\ 1, & \text{if } h \leq u \end{cases} \quad (4.19)$$

で記述されることは前章で述べた。このダイナミクスは、 $x(t)$ と $u(t)$ が同符号の場合、内部状態が大きくなるニューロンの影響を弱める効果をもつ。自己想起型の連想記憶では、記憶パターンが安定点であるため、記憶パターンで系の状態が停留する状況では $x(t)$ と $u(t)$ が同符号である。よって、このダイナミクスによって、内部状態を小さく保ち、大きい固有値をもつ固有ベクトルで張られる偽記憶への引き込みを防ぐことができる。しかしながら、 $x(t)$ と $u(t)$ が異符号であれば、その効果は全く逆になる。

荷重行列 W が非対称の場合、線形変換によって状態ベクトルが回転するので、安定なアトラクター上でも $x(t)$ と $u(t)$ が異符号になる場合がある。であるから、非対称結合系では、上で与えられた部分反転法を用いても、対称結合系と同様の効果は見られない。しかしながら、もし、 W の固有値の位相成分による回転効果を除いて、固有値の絶対値による拡大縮小のみを取り出し、絶対値の大きい固有値の影響を弱めることができれば、非対称結合系でも部分反転法やニューロウインドウ法を実現することができると考えられる。

一定の周期 Q をもつ時系列だけを記憶した相互相関連想記憶において、絶対値の大きい固有値をもつ固有ベクトル成分の拡大だけを抽出し、それを除去する想起ダイナミクスとして、部分反転法を

$$x_i(t+1) = \text{sgn} \left[\sum_j w_{ij} (x_j(t) - \lambda \phi(\tilde{u}_j(t))) \right] \quad (4.20)$$

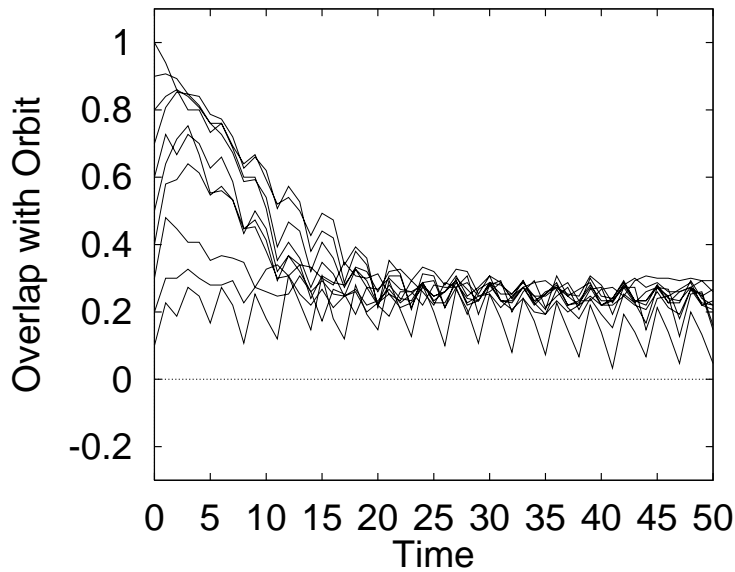
$$\tilde{u}_j(t) = \sum_{k=1}^N v_{jk} x_k \quad (4.21)$$

$$[v_{ij}] = W^Q \quad (4.22)$$

のように変形して使う方法が考えられる。ここで、関数 $\phi(u)$ は式 (4.19) で与えられた関数と同一である。周期が一定値 Q の相互相関連想記憶では、 Q ステップ後には最初のパターンに戻ってくるので、 Q ステップを 1 ステップと考えると自己連想記憶となる。その Q ステップの間に Q 回の W による変換を経るので、 W^Q は回転成分が除去されたほぼ対称の行列になっている。このことは、前節の固有空間解析において、周期 Q の系列を記憶した時の位相が $\frac{2n\pi}{Q}$ (n : 自然数) 付近になっていることから容易に予想される。よって、 W^Q をかけて内部状態が大きくなるニューロンは、大きい固有値を持つ空間の構成要素となっている。であるから、 W^Q を部分反転法に用いることで、記憶容量の向上ができると予想される。実際、相互相関連想記憶で記憶容量以上のパターンを記憶した場合の数値実験の結果を図 4.14 に示す。この図から、部分反転法に W^Q を用いることで、記憶容量が向上しているが分かる。

しかしながら、この W^Q を用いたダイナミクスは、同じ周期の系列を記憶した時にしか適用できない。また、たとえ覚えた系列の周期が一定でも、その周期が長くなると、記憶容量の向上幅は小さくなる。長い周期を記憶する場合に対応できない原因は以下のように説明される。前節の結果にあった通り、周期 Q の系列を記憶した場合は、固有値の位相が $\frac{2n\pi}{Q}$ (n : 自然数) 付近となるが、多少の誤差は含んでいる。これを Q 乗すると、位相の誤差が Q 倍になるので、位相のずれが大きくなる。その結果、 x と u の符号にずれが生じ、部分反転法が大きい固有値を持つ方向への流れを止める役割を果たせなくなってしまうわけである。

以上の制約に縛られず、より一般的に応用できる手法として、上記のダイナミクスにおける行列 $V = [v_{ij}]$ に $V = W^T W$ を利用する方法が考えられる。 W^T をかけると W と反対方向に位相を回転させるので、 W による回転をキャンセルする意味をもつ。この方法ならば、どの周期をもつ成分も同時に回転を抑制でき、異なる周期の系列を同時に記憶した場合にも有効に機能する。また、 V の計算に必要な掛け算は一回だけ済むため、多くの掛け算を繰り返すことによって生じる位相のずれの蓄積がなくなり、長い周期をもつ系列にも対応することができる。その結果、異なる周期や長い周期をもつ系列においても記憶容量の増大やニューロウインドウ法による選択的想起が可能になる。実際、相互相関連想記憶で記憶容量以上のパターンを記憶した場合の数値実験を図 4.15 に示す。この図のように、ここで提案した方法は、異なる長い周期の系列を記憶した場合でも、記憶容量の増加を実現することが分かる。また、図 4.16 に、通常ダイナミクス、 $W^T W$ を用いた部分反転法、 W^Q を用いた部分反転法を用いて同周期 Q の系列を記憶させたネットワークの想起する場合の記憶容量を比較した結果を示す。このように、 W^Q を用いた部分反転法では、 Q が大きくなるに従い記憶容量が減少するが、 $W^T W$ を用いた部分反転法では Q が増加しても $a = 0.4$ 近くの記憶容量を維持できることが分かる。



(a) 通常の想起過程

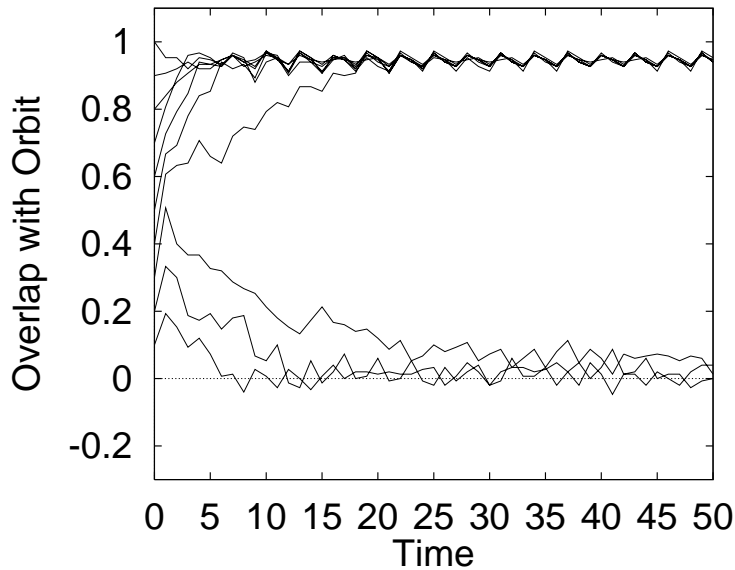
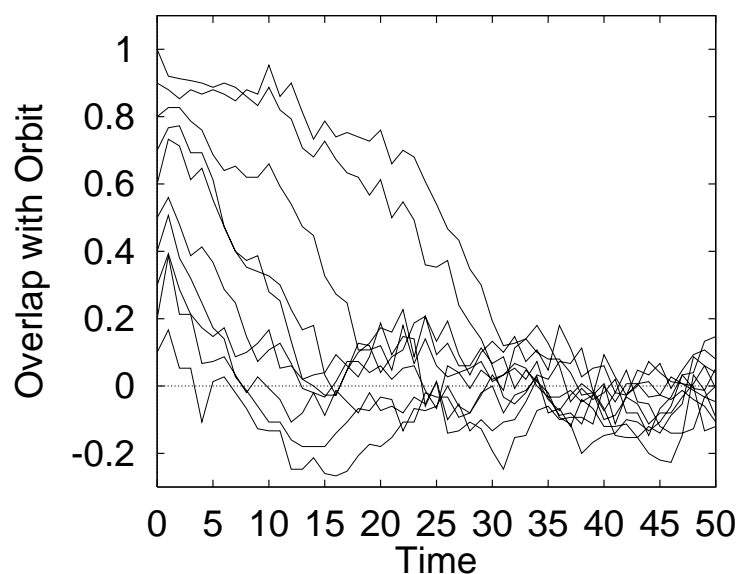
(b) W^Q を用いる部分反転法を利用した場合の想起過程 ($h = 1.5, \lambda = 0.5$)

図 4.14: 想起過程の数値実験 ($N = 300, P = 30, Q = 3$)。通常の想起法の記憶容量を超えた記憶率でも、 $V = W^Q$ を用いた部分反転法では想起に成功している。



(a) 通常の想起過程

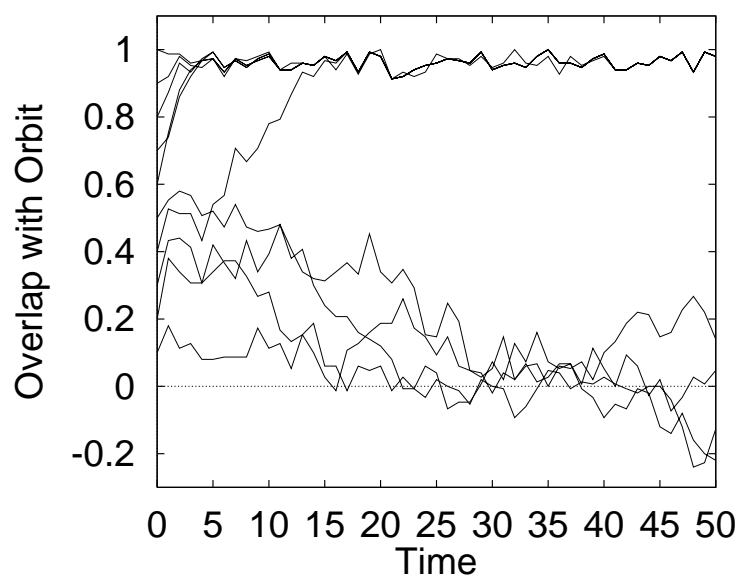
(b) $W^T W$ を用いる部分反転法を利用した場合の想起過程 ($h = 1.0, \lambda = 0.5$)

図 4.15: 想起過程の数値実験 ($N = 300, Q = 20, 30, 40$)。ここでは周期 $Q = 30$ のパターンの想起を試みている。通常の想起法の記憶容量を超えた記憶率でも、 $V = W^T W$ を用いた部分反転法では想起に成功している。

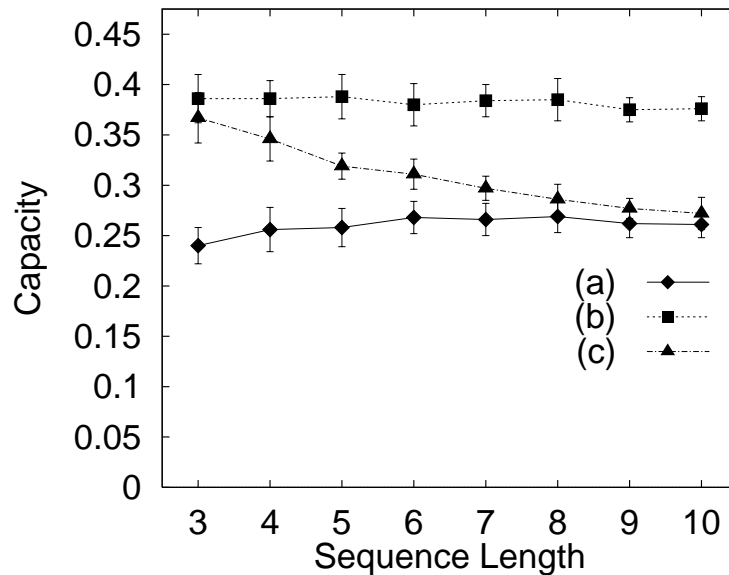


図 4.16: 同周期 Q の系列を相関学習で記憶したネットワークの記憶容量: (a) 通常の想起; (b) $W^T W$ を用いた部分反転法 ($h = 1.0, \lambda = 0.5$); (c) W^Q を用いた部分反転法 ($h = 0.5 \times Q, \lambda = 0.5$)。 W^Q を用いた部分反転法では、 Q が大きくなるに従い記憶容量が減少するが、 $W^T W$ を用いた部分反転法では Q が増加しても $a = 0.4$ 近くの記憶容量を維持することができる。

4.4 まとめ

本章では、非対称な結合荷重をもつニューラルネットワークについて、荷重行列の固有値と固有ベクトルを直接求めることで、そのダイナミクスの幾何構造を解析する方法を提案した。その解析法を用い、相互相関連想記憶、ランダムネットワークのダイナミクスを調べた。非対称の荷重を持つネットワークにおいても、対称のネットワークと同様に、固有値の絶対値が大きい空間へ状態遷移が進む。但し、非対称行列の固有値は一般に複素数になるので、長い周期をもつ状態がリミットサイクルとして生じることがある。

非対称行列の場合も、荷重行列 W の代わりに $W^T W$ を使う部分反転法を用いると、大きい固有値を持つ空間への状態遷移を防ぐことができる。これによって、相互相関連想記憶の記憶容量が向上することが示された。

第 5 章：ニューロウインドウ法

この章では、ニューラルネットワークを用いた連想記憶システムにおいて、新たな機能を実現するアルゴリズムとして、ニューロウインドウ法を提案する。まず、第1節でニューロウインドウ法の基本アルゴリズムを紹介する。続いて、第2節では自己連想記憶に、第3節では相互相関連想記憶にニューロウインドウ法を応用し、階層的な概念形成モデル等の新たな機能を持つ連想記憶システムの提案を行う。第4節においては、ニューロンの出力が $(0, 1)$ の二値で表現される場合について、ニューロウインドウ法の定式化を行う。最後に、第5節において、ニューロウインドウ法を用いて、オートマトンの状態遷移の効率化を学習するモデルを提案する。

5.1 ニューロウインドウ法

ニューロウインドウ法とは、第3章で紹介した部分反転法

$$x_i(t+1) = \text{sgn}\left(\sum_{j=1}^N w_{ij}(x_j - \lambda\phi(u_j(t)))\right) \quad (5.1)$$

$$u_j(t) = \sum_{k=1}^N w_{jk}x_k \quad (5.2)$$

$$\phi(u) = \begin{cases} -1, & \text{if } u < -h \\ 0, & \text{if } -h \leq u < h \\ 1, & \text{if } h \leq u \end{cases} \quad (5.3)$$

において、パラメータ h を可変パラメータとして使用方法である。第3章でも述べた通り、部分反転法は内部状態の絶対値が h より大きいニューロンの出力の影響を弱めることで、固有値の大きい固有ベクトルで張られる状態（偽記憶状態）への状態遷移を防いでいる。

ランダムパターンを一定の係数を用いて記憶する場合においては、記憶パターンの入力に対する各ニューロンの内部状態の絶対値は、どの記憶パターンでも同一の期待値を持つ。であるから、記憶パターンの想起において、記憶と偽記憶を分離するには、記憶状態の内部状態が持つ期待値と偽記憶状態の内部状態が持つ期待値の間にパラメータ h を固定しておけば十分であった。

しかしながら、より一般のパターンの記憶を考えると、上で述べたランダムパターンの仮定、及び学習係数一定の仮定は必ずしも成り立たない。記憶するパターンに相関構造が存在したり、学習係数に違いがある場合には、記憶するパターンによって、想起状

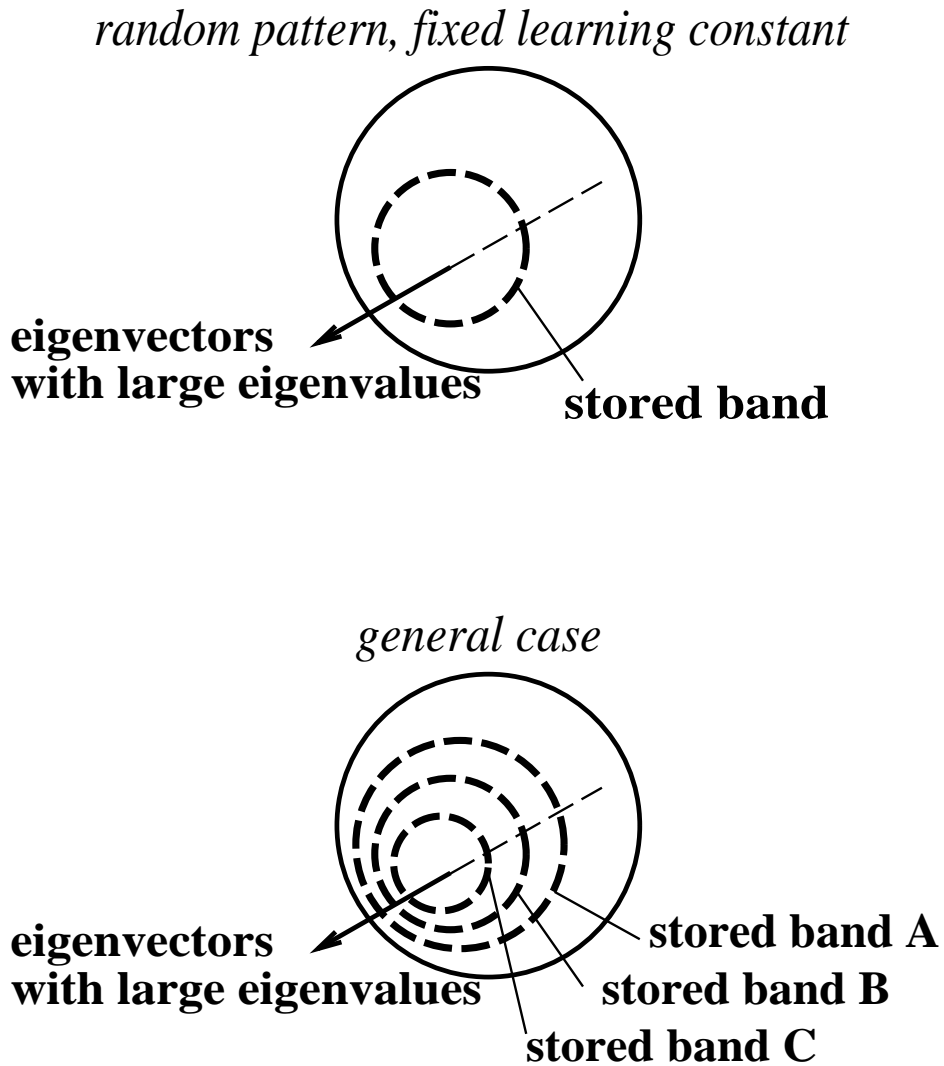


図 5.1: 一般の場合、ランダムパターンを同じ学習係数で記憶した自己相関連想記憶のように記憶パターンは一つのバンド (D 関数一定) にはおさまらず、多重のバンド構造が形成される。

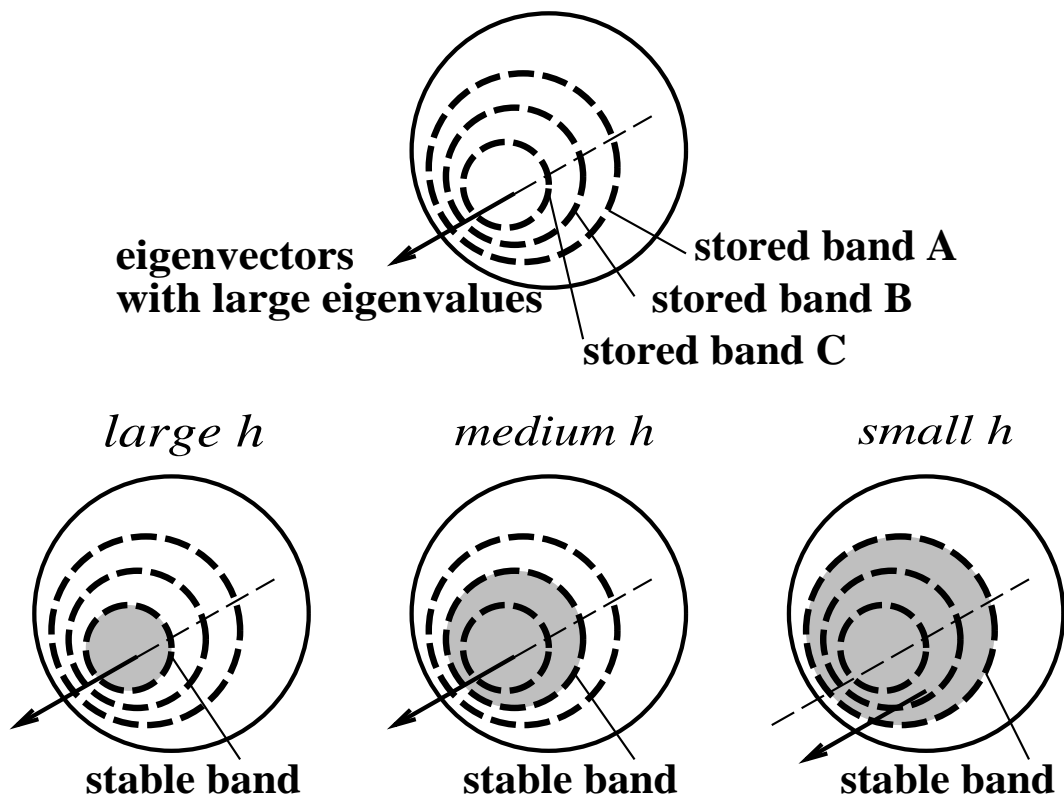


図 5.2: バンドが多重化された場合、部分反転法(非単調ニューロン)の閾値 h を制御することで各バンドを選択的に安定化できる。

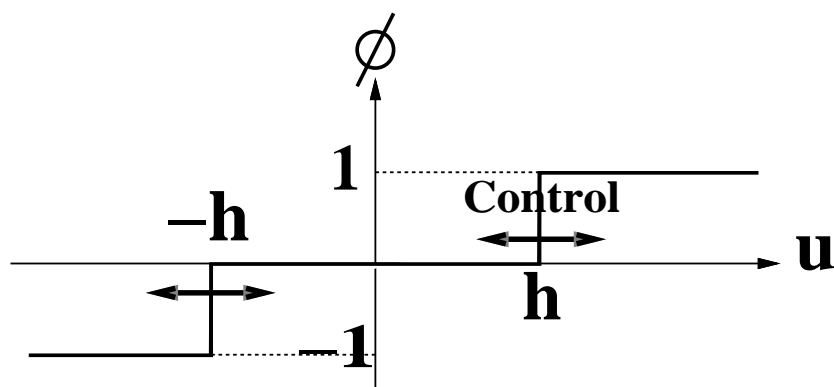


図 5.3: 離散時間ニューロウインドウ法の閾値制御。

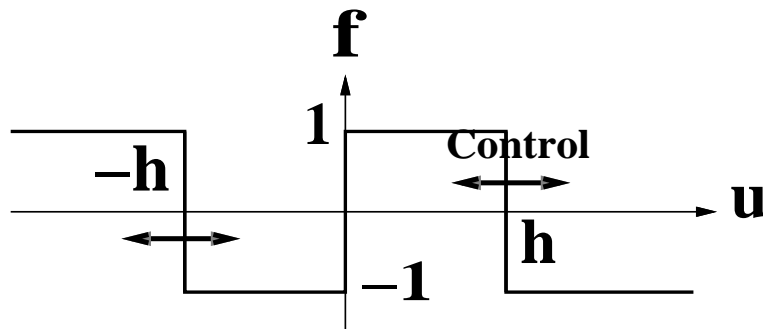


図 5.4: 連続時間ニューロウインドウ法の閾値制御。

態におけるニューロンの内部状態の期待値にばらつきが生じてくる。そのような場合では、記憶が構成する幾何構造も当然異なってくる。

ランダムパターンを一定の学習係数で記憶した場合と、より一般の場合とで、記憶空間の幾何構造がどのように異なってくるかを概略図として示したものが図5.1である。このように、一般の場合は、記憶パターンは状態ベクトルを張る超球面上の単一のバンドではなく、多くのバンド中に散らばって存在している。であるから、部分反転法である領域を不安定化すると、安定化されて想起可能になる記憶パターンと、逆に不安定化されてしまう記憶パターンが同時に発生する状況が生じることになる。

ここで、この状況を利用して、次のような応用を考えることができる。部分反転法において、パラメータ h を固定せず、状況によって任意に制御可能なパラメータとれば、 h を調節することで、多くの記憶パターンのバンドの中からあるバンドに含まれるパターン群をを 選択的に安定化し、それらを想起することができる(図5.2)。この方法では、図5.3に示す通り、関数 $\phi(u)$ の閾値 h を調節することで超球面上の安定領域を取り出している。これは、連続時間モデルでニューロンの活性化関数として図5.4のような窓型の関数を用い、その窓の大きさをパラメータ h を変えて想起を行うことに対応する(第7章参照)。この窓型関数の窓の大きさが、多重化されたバンド構造をなすパターン群の一つを取り出し安定化する作用をもつことから、この領域の選択的安定化法をニューロウインドウ法と呼ぶことにする。

次節以降では、このニューロウインドウ法を様々な連想記憶システムに応用することを試みることにする。

5.2 自己相関連想記憶のニューロウィンドウ法

自己相関連想記憶で、記憶パターンが多重バンドを形成する場合としては、パターンが異なる学習係数で記憶される場合と、相関構造をもつパターンが記憶される場合が挙げられる。以下の議論においては、それぞれのケースについて、ニューロウィンドウ法の適用を行う。

5.2.1 多様な学習係数を用いた連想記憶

この小節では、記憶バンドが多重化される例として、重みづけに変化をつけて記憶する場合について考察する。ここでは、最も単純な枠組として、 P 個 ($P = \text{偶数}$) の記憶パターンを半分に分け、二種類の重みづけで記憶する場合を考える。今、自己結合はないとすると、記憶行列は

$$w_{ij}^0 = \frac{B_1}{N} \sum_{\mu=1}^{\frac{P}{2}} s_i^\mu s_j^\mu + \frac{B_2}{N} \sum_{\mu=\frac{P}{2}+1}^P s_i^\mu s_j^\mu - \frac{(B_1 + B_2)P}{2N} \delta_{ij} \quad (5.4)$$

となる。ここで $B_1, B_2 (B_1 < B_2)$ は記憶に重みづけをするパターンの学習係数で、 δ_{ij} はクロネッカーのデルタである。

このようにして構成された記憶行列について、学習係数の比 $b = B_2/B_1$ が小さい場合 ($b = 2$) と大きい場合 ($b = 5$) の固有値の分布をそれぞれ図 5.5、図 5.6 に示す。

これらの図から分かるように、学習係数の比 b が小さい場合は、ランダムパターンを一定の係数で覚えた時の固有値分布を B_1 と B_2 を中心に二つ重ねあわせた形状の固有値分布が得られる。そして、 B_2/B_1 の値が大きくなると、記憶空間の固有値分布が二段にはっきり分かれる。

また、 b の値が小さい場合と大きい場合について、記憶パターンを固有ベクトルで展開した結果を図 5.7、図 5.8 にそれぞれ示す。この図から、小さい学習係数 B_1 で記憶したパターンは、固有値の小さな固有ベクトル成分をより多く持ち、大きい学習係数 B_2 で記憶したパターンは、固有値の大きな固有ベクトル成分をより多く持つことがわかる。つまり、学習係数の異なる記憶パターン間が直交に近い関係にある。そして、この分離は b が大きくなるほどはっきりする。このように、異なった学習係数を使ってパターンを記憶すると、記憶空間が分割されることになる。この構造は、第3章で述べた符号反転記憶法で得られた構造と同一である。

以上述べてきた複数の学習係数を用いた連想記憶が生み出す幾何構造を、状態ベクトルを張る超球面上にマップしたものを図 5.9 に示す。この図のように、線形変換のフロー上で、小さい学習係数 B_1 で記憶したパターン群は上流側に、大きい学習係数 B_2 で記憶したパターン群は下流側に位置することになる。よって、ニューロウィンドウ法で下流領域を不安定化すれば係数 B_1 で記憶したパターン群が安定化し、その下流領域の不安定化を解除すれば、状態遷移は下流の係数 B_2 で記憶したパターン群まで進むと予想される。

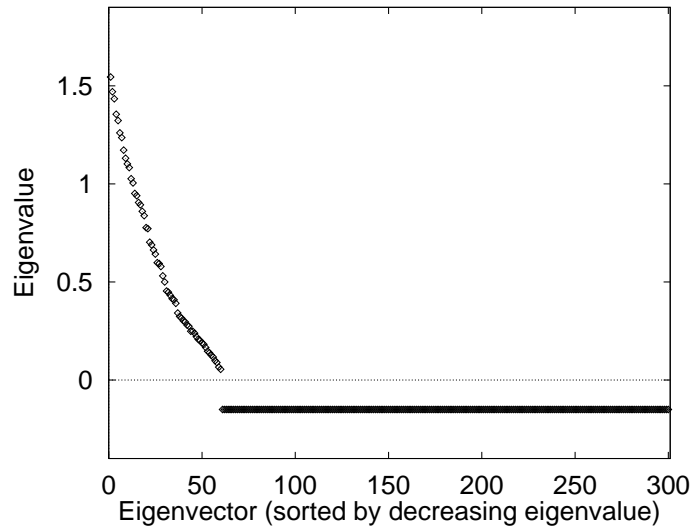


図 5.5: 学習係数 $B_1 = 0.5, B_2 = 1$ で $\frac{P}{2}$ 個ずつパターンを記憶した時の記憶行列 W^0 の固有値の分布 ($P = 60, N = 300$)。

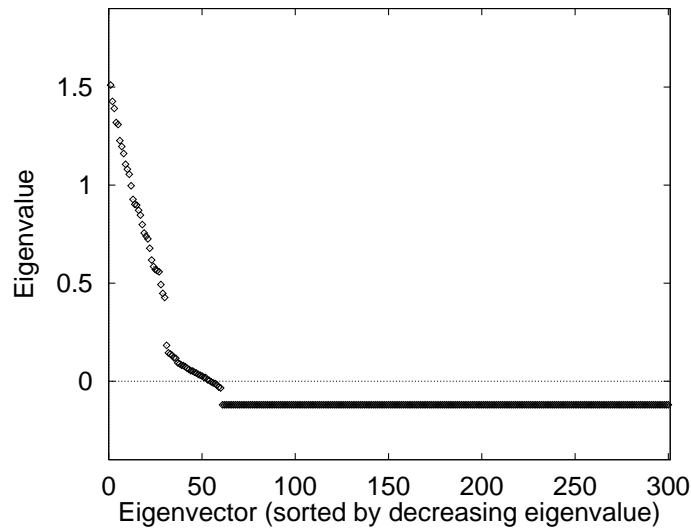
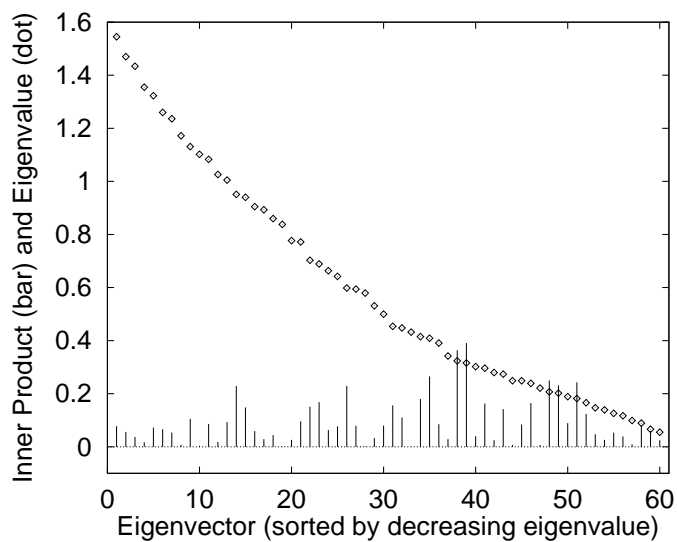
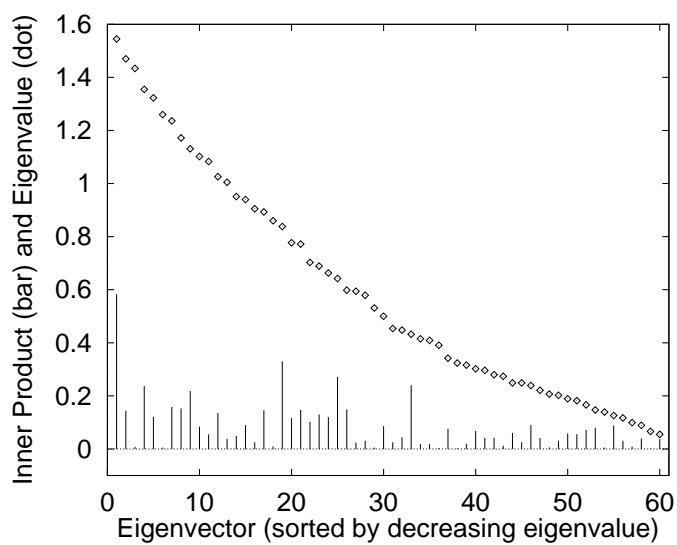


図 5.6: 学習係数 $B_1 = 0.2, B_2 = 1$ で $\frac{P}{2}$ 個ずつパターンを記憶した時の記憶行列 W^0 の固有値の分布 ($P = 60, N = 300$)。



(a) 学習係数 $B_1 = 0.5$ の記憶パターンと記憶行列の固有ベクトルの内積と固有値



(b) 学習係数 $B_2 = 1$ の記憶パターンと記憶行列の固有ベクトルの内積と固有値

図 5.7: 学習係数 $B_1 = 0.5, B_2 = 1$ で $\frac{P}{2}$ 個ずつパターンを記憶した記憶行列 W^0 の固有ベクトルで記憶パターンを展開した結果 ($P = 60, N = 300$)。

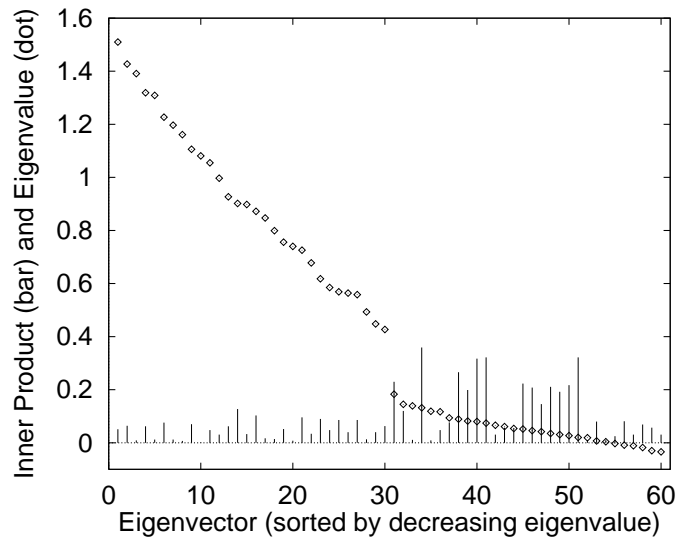
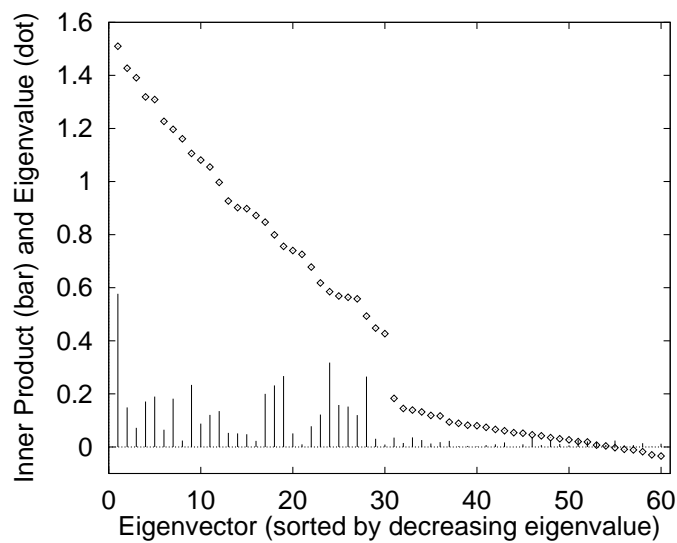
(a) 学習係数 $B_1 = 0.2$ の記憶パターンと記憶行列の固有ベクトルの内積と固有値(b) 学習係数 $B_2 = 1$ の記憶パターンと記憶行列の固有ベクトルの内積と固有値

図 5.8: 学習係数 $B_1 = 0.2, B_2 = 1$ で $\frac{P}{2}$ 個ずつパターンを記憶した記憶行列 W^0 の固有ベクトルで記憶パターンを展開した結果 ($P = 60, N = 300$)。

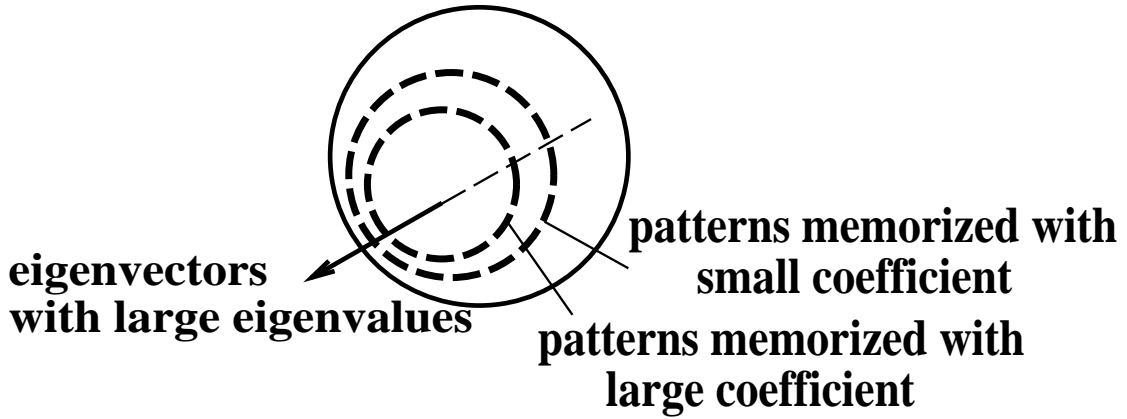


図 5.9: 異なる学習係数を用いた場合、小さい係数で記憶したパターンは上流にバンドを形成し、大きい係数で記憶したパターンは下流にバンドを形成する。

このようなメカニズムが実際の連想記憶で働き、学習係数が異なるパターン群の選択的想起が実現できるかどうかを数値実験によって確かめることにする。

ここでは、 $N = 1000$ 次元のベクトルで表現される 100 個のパターンをグループ1 とグループ2 の二つに 50 個ずつ分け、グループ1 のパターンを係数 $B_1 = 0.5$ で、グループ2 のパターンを係数 $B_2 = 1$ で覚えた場合について、想起過程におけるオーバーラップの変化を数値実験で調べた。図 5.10(a) は、ニューロウインドウ法で $h = 0.7$ とした想起過程におけるグループ1 のパターンとグループ2 のパターンとのオーバーラップの変化である。ここで、状態ベクトルの初期値は、着目しているグループ1 とグループ2 のパターンの中間的パターンとしている。この図から、ニューロウインドウ法で h を小さくすると、グループ1 のパターンが安定なアトラクタとして取り出せることが分かる。一方、図 5.10(b) は、ニューロウインドウ法で $h = 2.0$ とした想起過程におけるグループ1 とグループ2 に属するパターンとのオーバーラップの変化である。この図から、ニューロウインドウ法で h を大きくとった場合は、グループ2 のパターンが安定なアトラクタとして取り出せることが分かる。

この選択的想起における D 関数の変化を図 5.11 に示す。この図のように、グループ1 の想起では D 関数の増加が抑えられ、グループ2 の想起では D 関数が増加していく。つまり、 h を小さくすると、上流にあるグループ1 の記憶パターンを含むバンドが安定化し、 h を大きくすると、下流にあるグループ2 の記憶パターンを含むバンドが安定化されているわけである。更に系の変化を詳しく見るために、ニューロン数の次元が低い例で、各固有ベクトル成分の増減を調べた結果を図 5.12 に示す。この図から、 h の値が小

さいと、大きい固有値をもつ固有ベクトル成分の増加は抑えられるが、 h の値を大きくすると、大きい固有値をもつ固有ベクトルで張られる空間へ状態ベクトルが引き込まれていく様子がよく分かる。

以上の実験は、記憶率を固定しての話であった。しかし、記憶率を変化させると、連想記憶には記憶容量が存在することからも明らかなように、その系の性質は異なってくる。そこで、記憶率を変化させた時の記憶パターンの安定性と引き込み領域の大きさを数値実験により調べた結果を図5.13に示す。この図から、異なる記憶率における各グループのパターンの安定性をまとめると表5.1のようになる。この表から、この場合は記憶率が $a = 0.03$ から $a = 0.15$ の範囲で、ニューロウインドウ法を用いたグループ1とグループ2のパターンの選択的想起が可能となる。

ニューロウインドウ法を用いた選択的想起が実現できる記憶率の範囲は、学習係数の比によって変化する。学習係数の比 b が大きいと、大きい係数で記憶したパターンはより安定になり記憶容量は増すが、逆に小さい係数で記憶したパターンは不安定化し、ウインドウを用いても記憶容量が低下してしまう。よって、この場合はニューロウインドウ法を用いた選択的想起が機能する記憶率の範囲は非常に限られてしまう。一方、逆に学習係数の比 b が小さい場合は、小さい学習係数で記憶したパターン群も安定になるが、ウインドウのパラメータ h を大きくしても、両グループのパターンが共に安定化されるので、選択的想起が可能な範囲はやはり狭くなる。

以上、数値実験により、学習係数が異なる自己相関連想記憶において、ニューロウインドウ法の選択的想起が機能することを示した。しかし、このモデルに関する定量解析は示していないので、その理論的裏付け、および記憶容量や引き込み領域の理論値等の導出はここでは行わなかった。これらの定量的理論解析は第6章で述べる統計的手法を使って行うことが可能である。この統計手法を使った議論は次章に譲ることとし、次の小節では、記憶パターンに相関がある場合について解析を行うことにする。

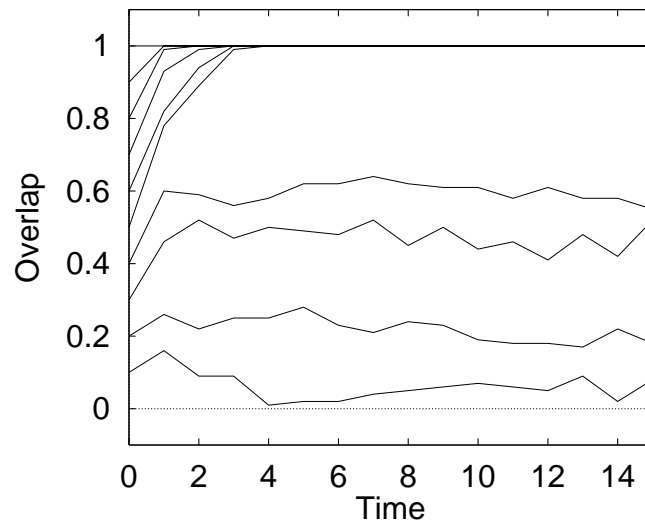
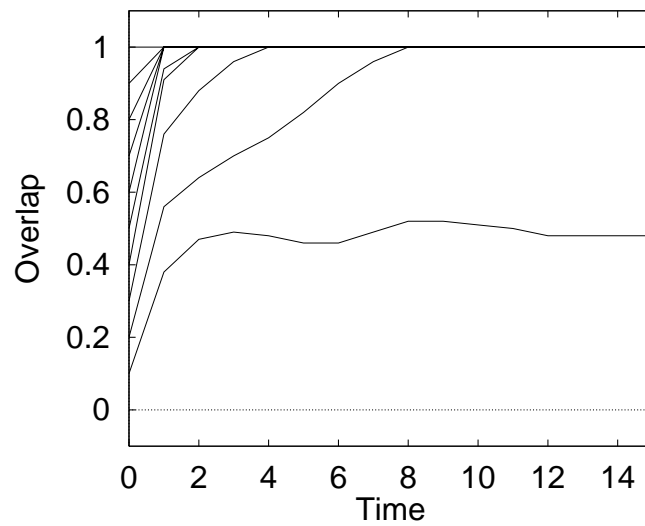
(a) グループ1 の想起 ($h = 0.7$)(b) グループ2 の想起 ($h = 2.0$)

図 5.10: ニューロウインドウ法を用いた想起過程におけるオーバーラップの変化 ($B_1 = 0.5$, $B_2 = 1.0$, $\lambda = 1.0$, $a = 0.1$, $N = 500$)。

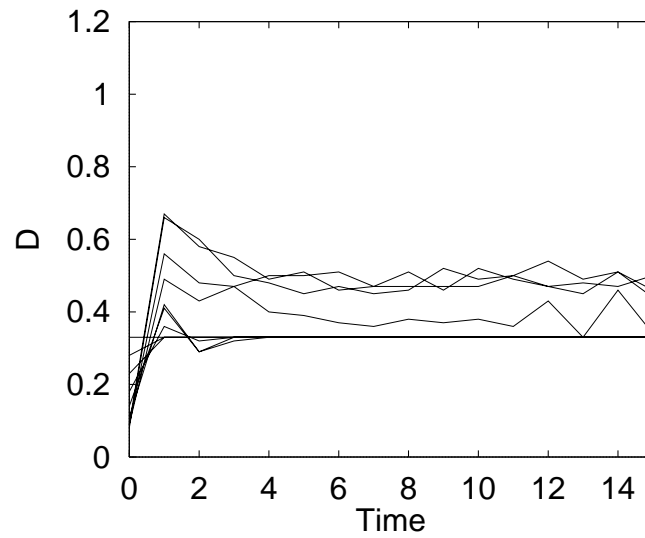
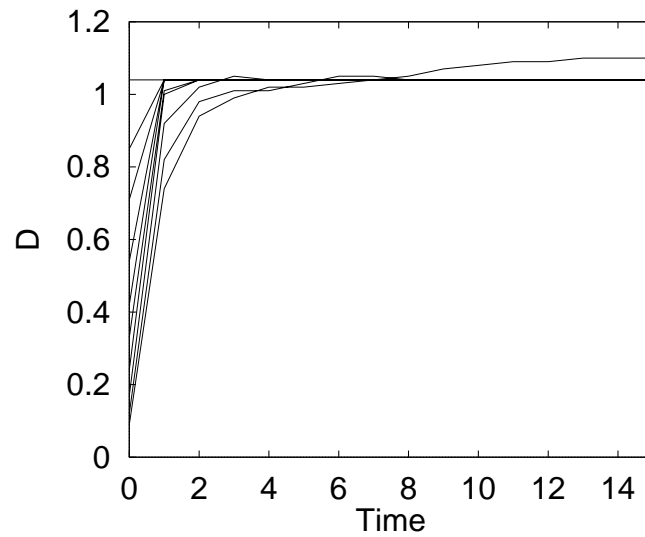
(a) グループ1 の想起 ($h = 0.7$)(b) グループ2 の想起 ($h = 2.0$)

図 5.11: ニューロウインドウ法を用いた想起過程における D 関数の変化 ($B_1 = 0.5$, $B_2 = 1.0$, $\lambda = 1.0$, $a = 0.1$, $N = 500$)。

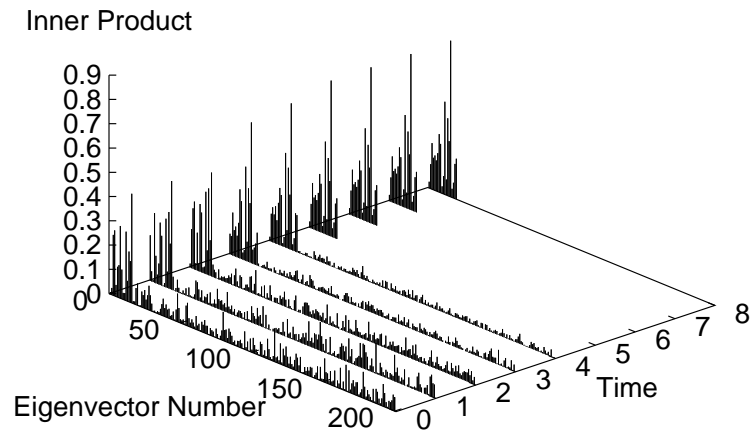
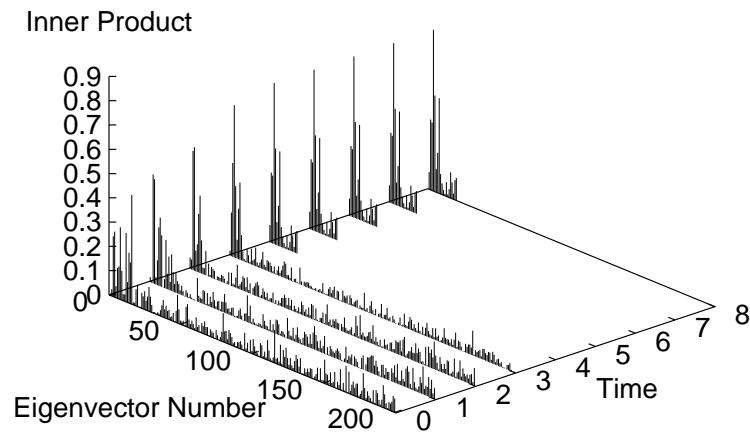
(a) グループ1 の想起 ($h = 0.75$)(b) グループ2 の想起 ($h = 2.0$)

図 5.12: ニューロウィンドウ法を用いた想起過程における全固有ベクトル成分の変化 ($B_1 = 0.5$, $B_2 = 1.0$, $\lambda = 1.0$, $a = 0.1$, $N = 200$)。

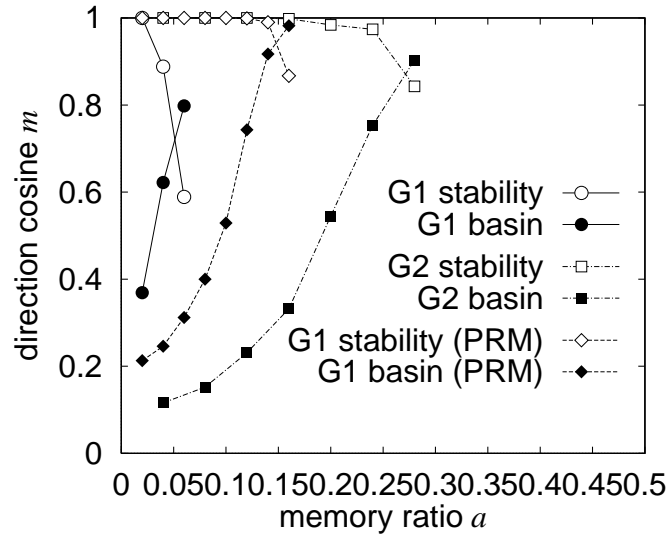


図 5.13: 二種類の学習係数を用いて記憶した自己相関連想記憶における記憶パターンの安定性と引き込み領域。学習係数 0.5 を用いて記憶したグループ 1(G1) を通常のダイナミクスと部分反転法 (PRM: $h = 1.5$, $\lambda = 1.0$) のダイナミクスで想起した場合、及び学習係数 1.0 を用いて記憶したグループ 2(G2) を通常のダイナミクスで想起した場合の結果を示した ($P_1 = P_2$, $N = 500$)。

記憶率	group1 ($h = 1.5$)	group1 ($h = \infty$)	group2 ($h = 1.5$)	group2 ($h = \infty$)
0.00-0.03	安定	安定	不安定	安定
0.03-0.15	安定	不安定	不安定	安定
0.15-0.25	不安定	不安定	不安定	安定
0.25-	不安定	不安定	不安定	不安定

表 5.1: 数値実験の結果から評価した各記憶率における記憶パターンの安定性。

5.2.2 相関をもつパターンの連想記憶

前小節では、異なった学習係数を使って記憶する場合についてニューロウインドウ法を導入した。本小節と次の小節では、ニューロウインドウ法が応用できるもう一つのケースとして、相関のあるパターンを記憶する連想記憶モデルを取り上げることにする。中でも、この小節では、単純な相関構造をなす記憶パターン群を考える。

以下の議論では、相関構造をなす記憶パターン群として、 P_1 個のランダムパターンを親パターンとして、それぞれの親パターンとの相関が r の子パターンを P_2 個ずつ生成し、それらを記憶パターンとして記憶する場合を考える。今、 μ 番目の親パターンに由来する ν 番目の子パターンとして生成された記憶パターンを $s^{(\mu,\nu)}$ と記述すると、パターン間の相関構造は

$$\frac{1}{N} E[s^{(\mu,\nu)} \cdot s^{(\mu',\nu')}] = \begin{cases} 1, & \text{if } \mu = \mu', \nu = \nu' \\ R(= r^2), & \text{if } \mu = \mu', \nu \neq \nu' \\ 0, & \text{if } \mu \neq \mu' \end{cases} \quad (5.5)$$

となる。このようなクラスター型相関構造をもつパターンを考えた時、各クラスターの中心に位置するクラスターの代表パターンとして、

$$s_i^{(\mu)} = \text{sgn}\left(\sum_{\nu=1}^{P_2} s_i^{(\mu,\nu)}\right) \quad (5.6)$$

で表される概念パターン $s^{(\mu)}$ を定義することができる [2]。

このようなクラスター型の相関をなすパターンを自己相関学習で記憶すると、荷重行列 (自己結合なし) は

$$w_{ij}^0 = \frac{1}{N} \sum_{\mu=1}^{P_1} \sum_{\nu=1}^{P_2} s_i^{(\mu,\nu)} s_j^{(\mu,\nu)} - a\delta_{ij} \quad (5.7)$$

で与えられることになる。但し、ここで a は記憶率 $a = \frac{P_1 \times P_2}{N}$ である

ここで与えられた荷重行列の固有値の分布を図 5.14 に示す。図 5.14 から、相関のあるパターンを記憶した場合も、二種類の学習係数を用いて記憶した場合と同様に、記憶空間の固有値の分布が二段になることが分かる。次に、この荷重行列の固有ベクトルと記憶パターンベクトル、および概念パターンベクトルの内積を計算した結果を図 5.15 に示す。これらの図から分かるように、概念パターンは記憶パターンよりも大きい固有値をもつ固有ベクトル群の成分を多く持つ。このことから、線形変換のフローにおいて、概念パターンは記憶パターンよりも下流に位置していることが分かる。よって、記憶パターン想起時における内部状態の絶対値の期待値

$$E[|u_i^{(\mu,\nu)}|] = E\left[\left|\sum_{j=1}^N w_{ij} s_j^{(\mu,\nu)}\right|\right] \quad (5.8)$$

と、概念パターン想起時の内部状態の絶対値の期待値

$$E[|u_i^{(\mu)}|] = E\left[\left|\sum_{j=1}^N w_{ij} s_j^{(\mu)}\right|\right] \quad (5.9)$$

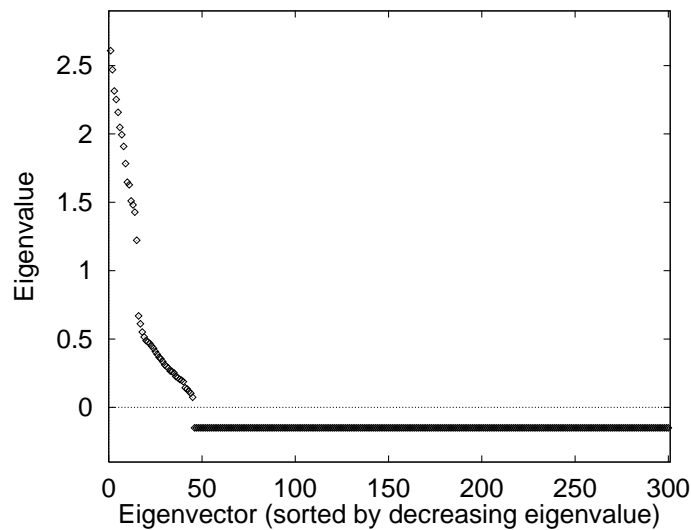


図 5.14: 相関のあるパターンを記憶した自己相関連想記憶における荷重行列の固有値の分布 ($P_1 = 15, P_2 = 3, R = 0.49, N = 300$)。

の間には

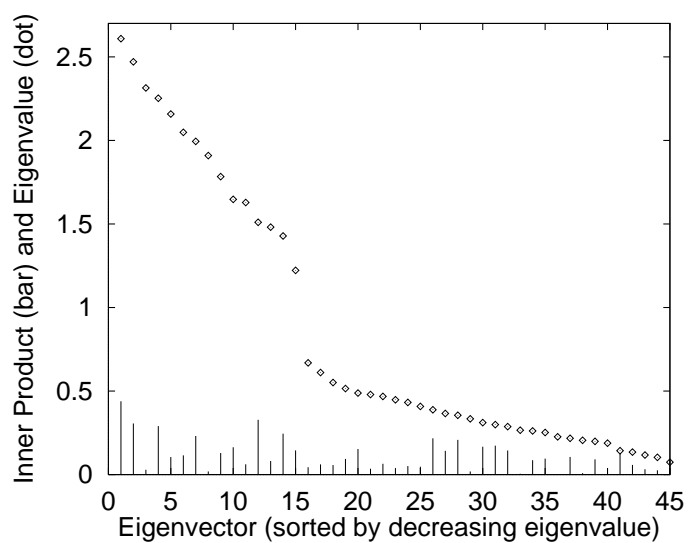
$$E[\|u_i^{(\mu, \nu)}\|] < E[\|u_i^{(\mu)}\|] \quad (5.10)$$

の関係が成り立っている。但し、この関係が成り立つのはクラスター内の相関がある程度強い場合に限られる。相関の違いによる記憶パターン、概念パターンの D 関数の値の変化をプロットしたものを図 5.16 に示す。この図から、 $P_2 = 3$ の場合、概念パターンが下流に位置するのは $r > 0.3 (R > 0.09)$ の範囲であることが分かる。

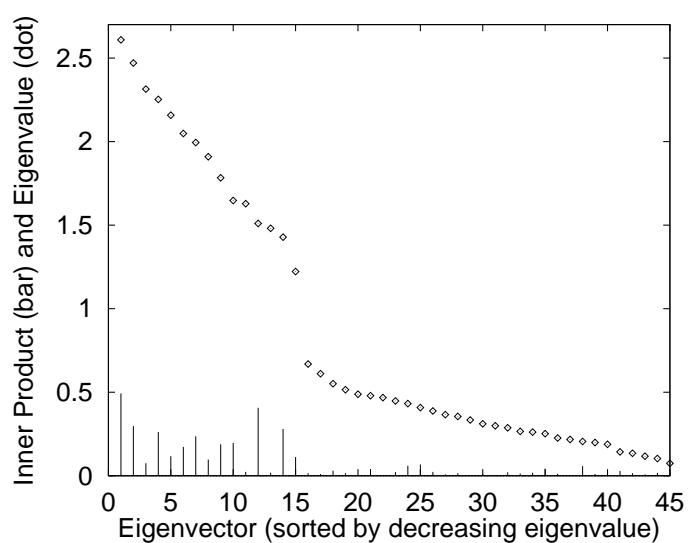
以上の幾何構造から、ニューロウインドウ法を利用することで上流に位置する記憶パターンと下流に位置する概念パターンを図 5.17 のように選択的に想起することができる。実際、 $N = 1000, P_1 = 10, P_2 = 3, R = 0.49$ の例において、ウインドウのパラメータ h を変化させ、想起過程におけるオーバーラップを数値実験により求めた結果を図 5.18 に示す。このように、パラメータ h が小さい場合は記憶パターンの想起が、パラメータ h が大きい場合は概念パターンの想起が実現されていることが分かる。

この想起過程における D 関数と固有ベクトル成分の変化を図 5.19 と図 5.20 に示す。これらの図から、想起過程においては、固有値の大きな固有ベクトル成分が徐々に増加するが、値の小さな h を用いる場合は記憶パターンで状態遷移が停止し、値の大きな h を用いる場合は概念パターンまで状態遷移が進んでいることが分かる。

上の数値実験は、記憶率が 0.1 の一例をとってニューロウインドウ法が機能することを示した。次に、ニューロウインドウ法が機能する記憶率の範囲について数値解析を行う。図 5.21 にニューロウインドウ法を用いたシミュレーションで得られた、各記憶率に



(a) 記憶パターンベクトルと固有ベクトルの内積



(b) 概念パターンベクトルと固有ベクトルの内積

図 5.15: 相関パターンを記憶した荷重行列の固有空間における記憶パターンと概念パターンの表現 ($R = 0.49$, $P_1 = 15$, $P_2 = 3$, $N = 300$)。概念パターンは記憶空間の中でも大きい固有値を持つ固有ベクトル群の成分を多く持つ。

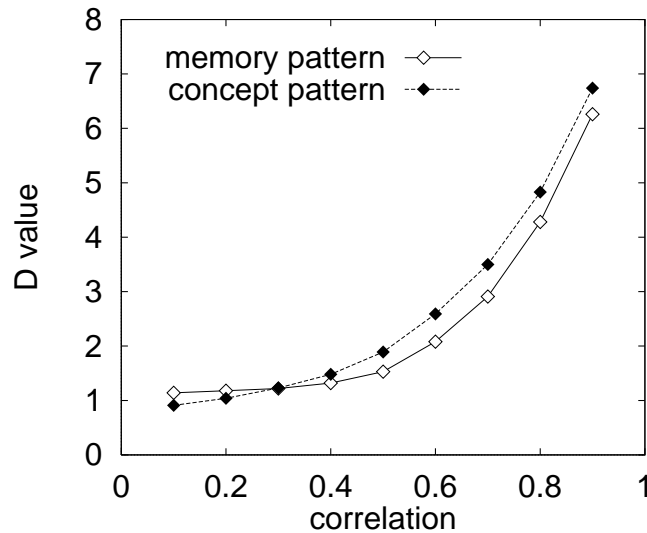


図 5.16: 記憶パターン、概念パターン入力時における D 関数の値 ($P_2 = 3$)。 $r > 0.3$ の領域では概念パターンの方が線形フローの下流に位置する。

における記憶パターン、概念パターンの安定性と引き込み領域を示す。この場合、記憶率が 0.15 までは選択的想起が可能になっていることが分かる。

図 5.21 で得た記憶容量は、式 (5.3) で定義される閾値関数をニューロウインドウ法に用いて得られる記憶容量である。ここで、ニューロウインドウ法で使用する関数 $\phi(u)$ に

$$\phi(u) = \gamma u \quad (5.11)$$

を用いると、記憶容量は図 5.22 のように増加する。このように、線形な関数をウインドウ法に用いることは、荷重行列を

$$V = W - \gamma W^2 \quad (5.12)$$

と変形することと等価である。この変形された行列 V で D 関数 $D(\mathbf{x}) = \mathbf{x}V^T V \mathbf{x}$ を考えると、 γ が大きくなると、記憶パターンと概念パターンの位置関係が逆転し、記憶パターンが線形変換で与えられるフローの下流に位置するようになることが分かる。その結果、記憶パターンが概念パターンより安定化されることになる。

以上の議論は $R = 0.49$ に固定しての結果であったが、この記憶容量は相関 R によっても異なる。また、相関 R が異なれば、選択的想起に適切なパラメータ h の大きさも異なる。各相関 R による記憶容量、および適切なパラメータ h 等の解析は次章の統計的議論において行うことにする。

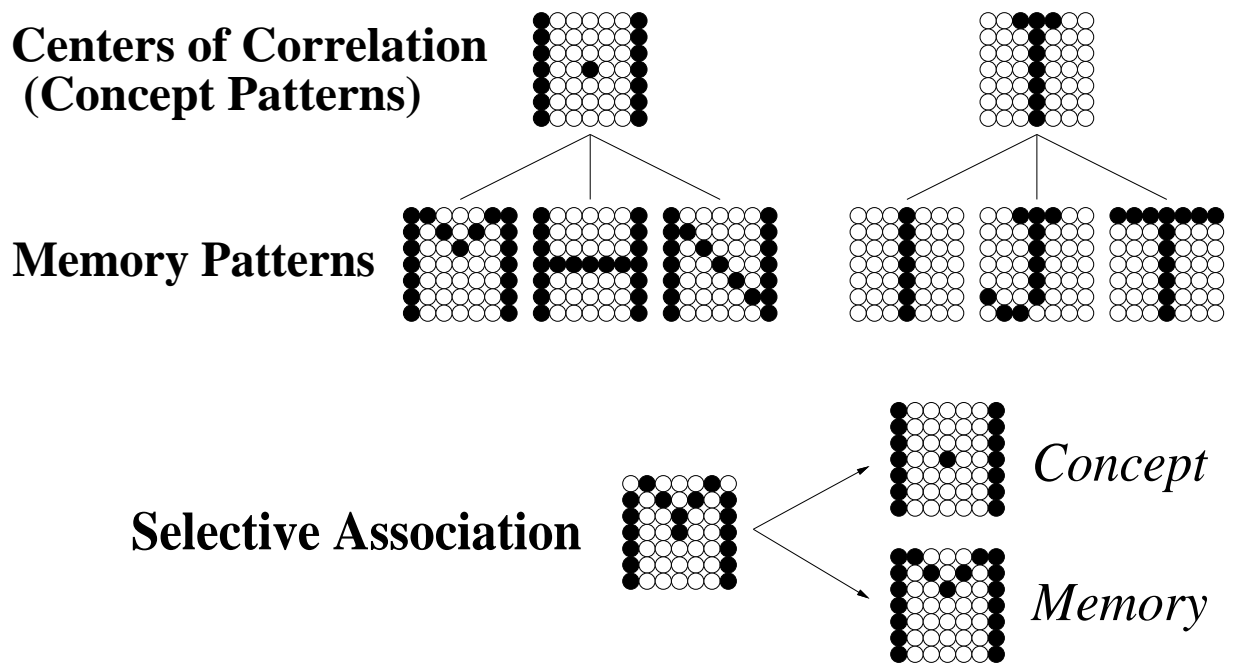


図 5.17: ニューロウインドウ法による選択的想起。

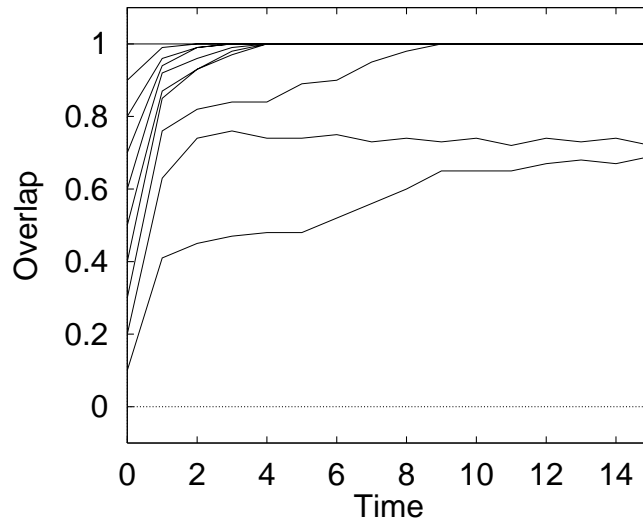
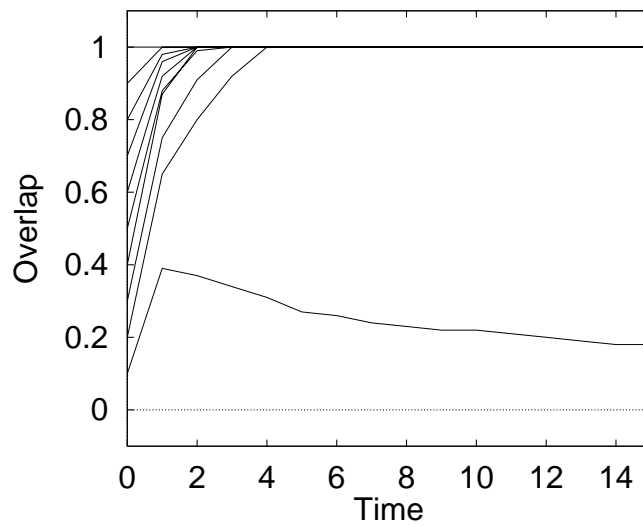
(a) 記憶パターンの想起 ($h = 1.5$)(b) 概念パターンの想起 ($h = 2.5$)

図 5.18: ニューロウインドウ法による選択的想起過程におけるオーバーラップの変化 ($N = 1000$, $P_1 = 25$, $P_2 = 3$, $R = 0.49$, $\lambda = 0.8$)。

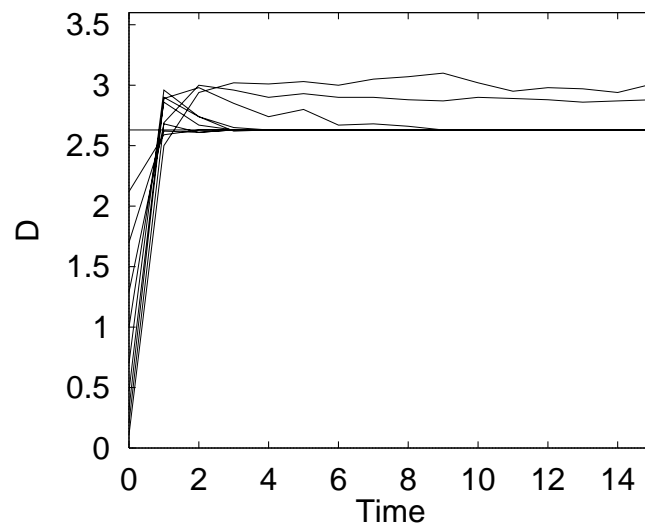
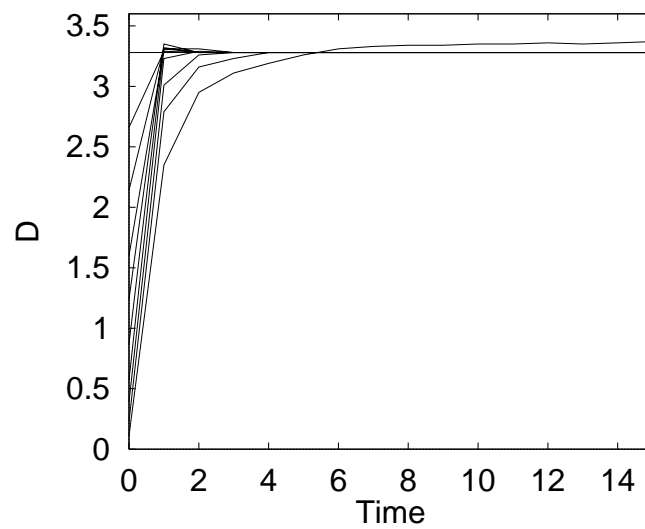
(a) 記憶パターンの想起 ($h = 1.5$)(b) 概念パターンの想起 ($h = 2.5$)

図 5.19: ニューロウインドウ法による選択的想起過程における D 関数の変化 ($N = 1000$, $P_1 = 25$, $P_2 = 3$, $R = 0.49$, $\lambda = 0.8$)。

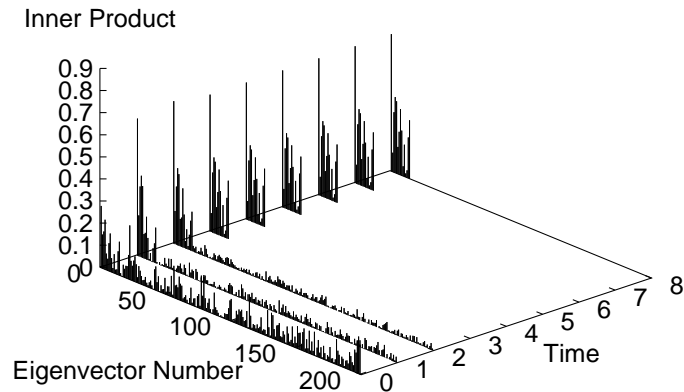
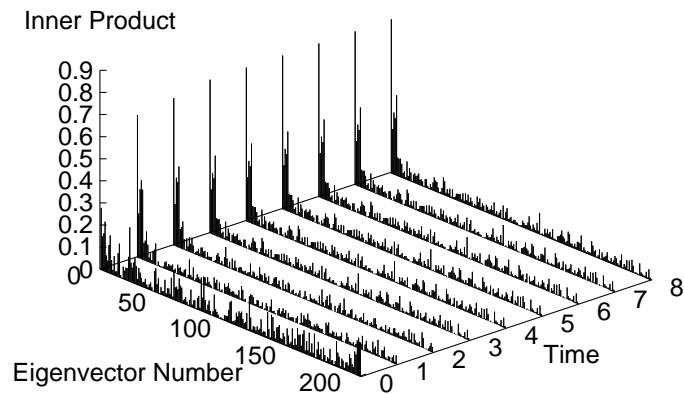
(a) 記憶パターンの想起 ($h = 1.5$)(b) 概念パターンの想起 ($h = 2.5$)

図 5.20: ニューロウインドウ法による選択的想起過程における全固有ベクトル成分の変化 ($N = 200$, $P_1 = 5$, $P_2 = 3$, $R = 0.49$, $\lambda = 0.8$)。

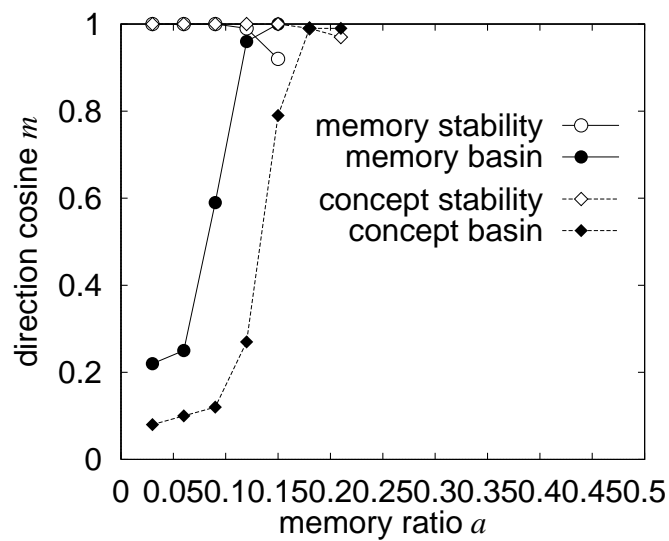


図 5.21: ニューロウインドウ法を用いたときの記憶パターン ($h = 1.5$, $\lambda = 0.8$) と概念パターン ($h = 2.5$, $\lambda = 0.8$) の安定性と引き込み領域 ($R = 0.49$, $P_2 = 3$, $N = 1000$)。

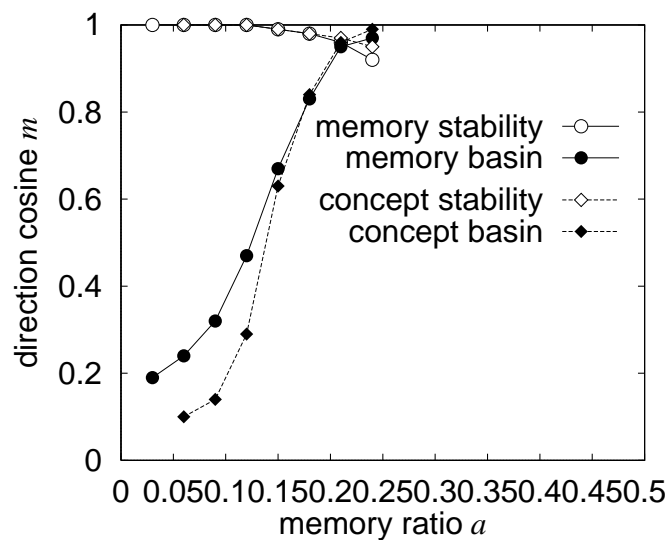


図 5.22: 線形関数を使ったニューロウインドウ法における記憶パターン ($\gamma = 0.3$) と概念パターン ($\gamma = 0.2$) の安定性と引き込み領域 ($R = 0.49$, $P_2 = 3$, $N = 1000$)。

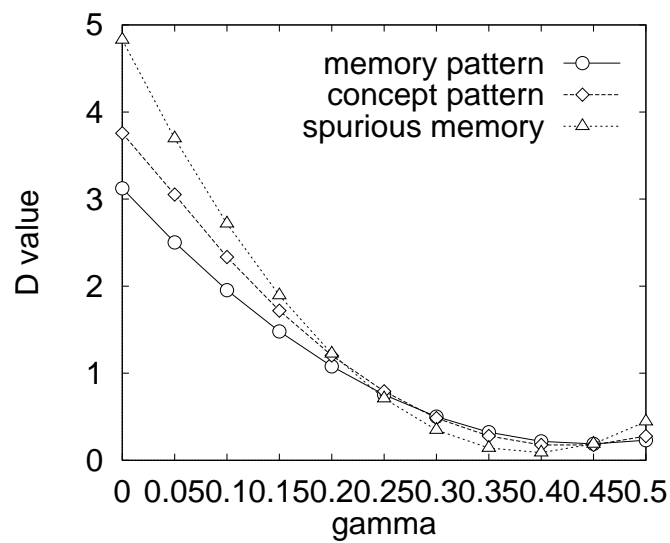


図 5.23: 荷重行列 $V = W - \gamma W^2$ で定義される D 関数 ($R = 0.49$, $P_1 = 60$, $P_2 = 3$, $N = 1000$) で見る記憶パターンと概念パターンの位置関係。 γ が大きくなると記憶パターンが線形フローの最下流に位置するようになる。

5.2.3 階層的相関をもつパターンの連想記憶

前小節で議論した記憶パターンと概念パターンの選択的想起モデルは、パターンの相関構造が三層以上の場合にも拡張できる。

以下の議論においては、多層の相関構造をもつパターン群として

$$\frac{1}{N} E[\mathbf{s}^{(p_1, p_2, \dots, p_L)} \cdot \mathbf{s}^{(q_1, q_2, \dots, q_L)}] = \begin{cases} 1, & \text{if } p_1 = q_1, \dots, p_L = q_L \\ R^{L-l}, & \text{if } p_1 = q_1, \dots, p_l = q_l, p_{l+1} \neq q_{l+1} \ (1 \leq l \leq L-1) \\ 0, & \text{if } p_1 \neq q_1 \end{cases} \quad (5.13)$$

で表される超計量的 (ultrametric) 階層的相関構造を持つパターン (図 5.24) を記憶する場合を考える。この場合、概念パターンも多層にわたって定義される。例えば、第 l 層の概念パターン $\mathbf{s}^{(p_1, p_2, \dots, p_l)}$ の成分は

$$s_i^{(p_1, p_2, \dots, p_l)} = \text{sgn} \left(\sum_{p_{l+1}=1}^{P_{l+1}} \sum_{p_{l+2}=1}^{P_{l+2}} \dots \sum_{p_L=1}^{P_L} s_i^{(p_1, p_2, \dots, p_L)} \right) \quad (5.14)$$

で与えられることになる。つまり、概念パターンとして $\mathbf{s}^{(p_1)}, \mathbf{s}^{(p_1, p_2)}, \dots, \mathbf{s}^{(p_1, p_2, \dots, p_{L-1})}$ が与えられる ($\mathbf{s}^{(p_1, p_2, \dots, p_L)}$ が記憶パターン)。

このような多層の相関構造を持つパターンを自己相関連想記憶で記憶する時、その荷重行列は

$$w_{ij} = \frac{1}{N} \sum_{p_1=1}^{P_1} \sum_{p_2=1}^{P_2} \dots \sum_{p_L=1}^{P_L} s_i^{(p_1, p_2, \dots, p_L)} s_j^{(p_1, p_2, \dots, p_L)} - a \delta_{ij} \quad (5.15)$$

$$a = \frac{1}{N} \prod_{l=1}^L P_l \quad (5.16)$$

で与えられる。

以下の議論では、相関が三層の場合について、荷重行列の固有空間解析を調べる。この時は、記憶パターン $\mathbf{s}^{(p_1, p_2, p_3)}$ と第一概念パターン $\mathbf{s}^{(p_1)}$ 、第二概念パターン $\mathbf{s}^{(p_1, p_2)}$ の二種類の概念パターンが定義される。まず、荷重行列の固有値の分布を図 5.25 に示す。図 5.25 から、三層の相関を持つパターンを覚えた自己相関連想記憶では、記憶空間の固有値の分布が三段になることが分かる。次に、この荷重行列の固有ベクトルと記憶パターンベクトル、および概念パターンベクトルの内積を計算した結果を図 5.26 に示す。これらの図から分かるように、記憶パターンは主に上位 $P_1 \times P_2 \times P_3$ 個の固有値をもつ固有空間 (記憶空間) で張られ、第二概念パターンは主に上位 $P_1 \times P_2$ 個の固有値をもつ固有ベクトル群で張られ、第一概念パターンは主に最も大きい P_1 個の固有値をもつ固有ベクトル群で張られている。このことから、線形変換のフローにおいて、記憶パターンが上流、第二概念パターンが下流、第一概念パターンが最下流に位置していることが分かる。よって、記憶パターン想起時における内部状態の絶対値の期待値

$$E[|u_i^{(p_1, p_2, p_3)}|] = E\left[\left|\sum_{j=1}^N w_{ij} s_j^{(p_1, p_2, p_3)}\right|\right] \quad (5.17)$$

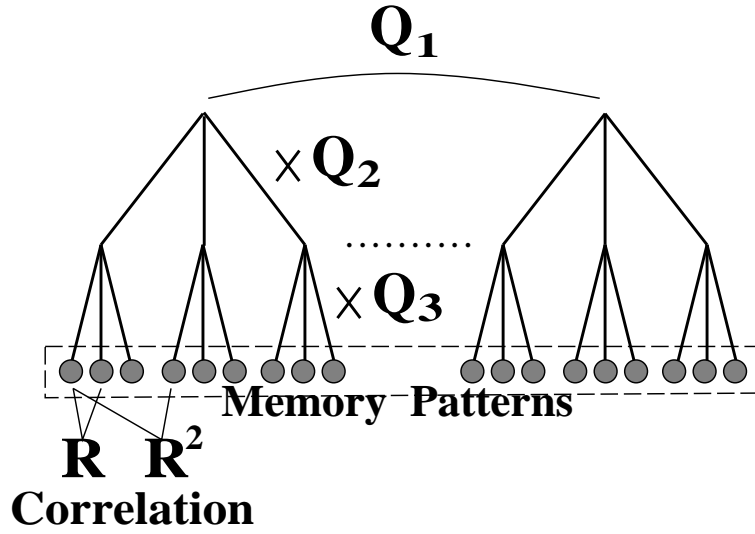


図 5.24: 超計量的 (ultrametric) 階層的相関構造 ($L = 3, P_2 = P_3 = 3$)。ノードが概念パターンに対応する。

第二概念パターン想起時における内部状態の絶対値の期待値

$$E[|u_i^{(p_1, p_2)}|] = E\left[\left|\sum_{j=1}^N w_{ij} s_j^{(p_1, p_2)}\right|\right] \quad (5.18)$$

第一概念パターン想起時における内部状態の絶対値の期待値

$$E[|u_i^{(p_1)}|] = E\left[\left|\sum_{j=1}^N w_{ij} s_j^{(p_1)}\right|\right] \quad (5.19)$$

の間には

$$E[|u_i^{(p_1, p_2, p_3)}|] < E[|u_i^{(p_1, p_2)}|] < E[|u_i^{(p_1)}|] \quad (5.20)$$

の関係が成り立っている。

このような幾何構造をもつ場合は、前小節の議論と同様に、ニューロウインドウ法を用いた選択的想起が可能になると予想される。実際、 $N = 1000, P_1 = 10, P_2 = 3, R = 0.49$ の例において、ウィンドウのパラメータ h を変化させ、想起過程におけるオーバーラップを数値実験により求めた結果を図 5.27 に示す。このように、パラメータ h が小さい場合は記憶パターンの想起が実現され、パラメータ h を大きくすると第二概念パターンの想起が、更に h を大きくすると第一概念パターンの想起が実現されていることが分かる。

この想起過程における D 関数と固有ベクトル成分の変化を図 5.28、図 5.29 に示す。この図のように、想起過程においては、固有値の大きな固有ベクトル成分が徐々に増加す

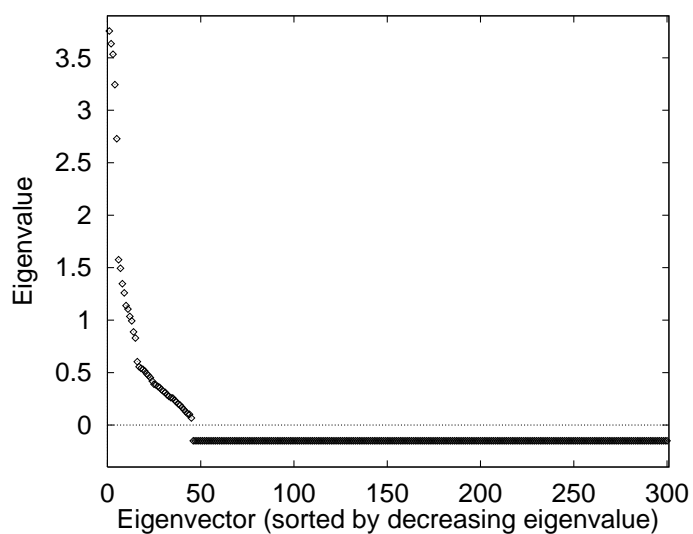
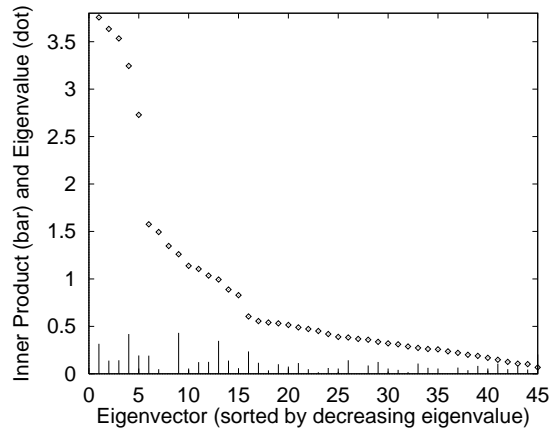


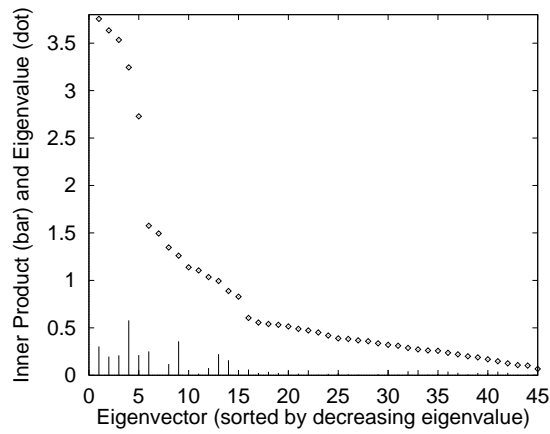
図 5.25: 階層的相関のあるパターンを記憶した自己相関連想記憶における荷重行列の固有値の分布 ($P_1 = 5, P_2 = 3, P_3 = 3, R = 0.49, N = 300$)。

るが、値の小さな h を用いる場合は記憶パターンで状態遷移が停止し、 h の値を大きくしていくと、第二概念パターンまで状態遷移が進み、更に h を大きくすると第一概念パターンまで状態遷移が進んでいくことが分かる。

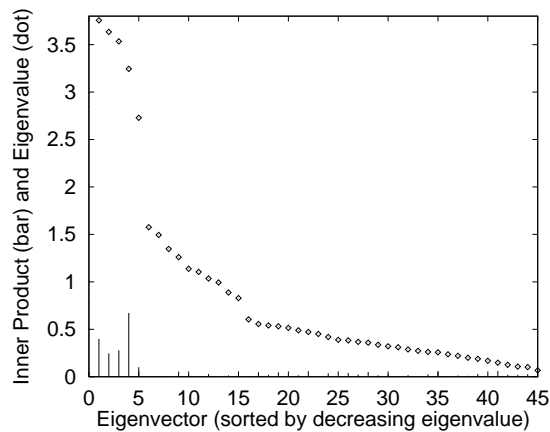
本小節で議論した階層的な概念形成と選択的想起の簡単な例を図 5.30 に示す。このように、ニューロウインドウ法を用いることにより、多層の相関クラスターをなすパターンを記憶した場合、パターンの分類の細かさを任意に変化させ想起を行うことが可能となった。



(a) 記憶パターンベクトルと固有ベクトルの内積



(b) 第二概念パターンベクトルと固有ベクトルの内積



(c) 第一概念パターンベクトルと固有ベクトルの内積

図 5.26: 階層的相関パターンを記憶した荷重行列の固有空間における記憶パターンと概念パターンの表現 ($R = 0.49$, $P_1 = 5$, $P_2 = P_3 = 3$, $N = 300$)。概念パターンはより大きい固有値を持つ固有ベクトル成分を多く持つ。

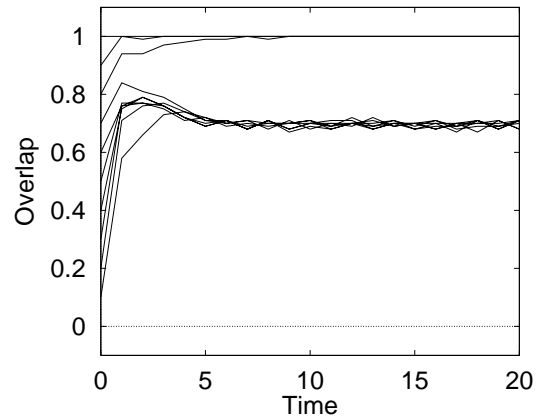
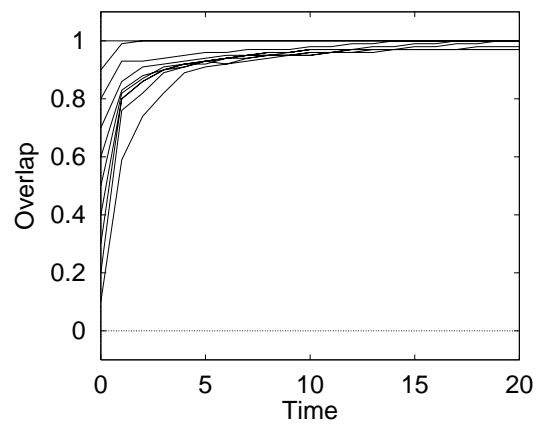
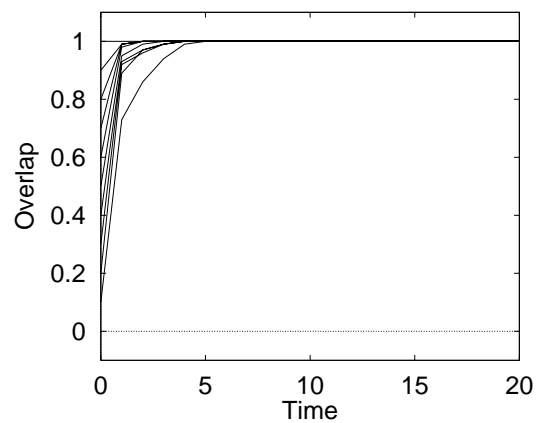
(a) 記憶パターンの想起 ($h = 2.5$)(b) 第二概念パターンの想起 ($h = 2.9$)(c) 第一概念パターンの想起 ($h = 5.0$)

図 5.27: ニューロウインドウ法による選択的想起過程におけるオーバーラップの変化 ($N = 1000$, $P_1 = P_2 = P_3 = 3$, $R = 0.49$, $\lambda = 1.0$)。

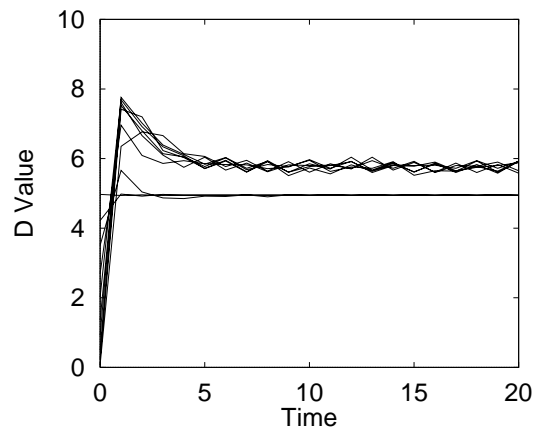
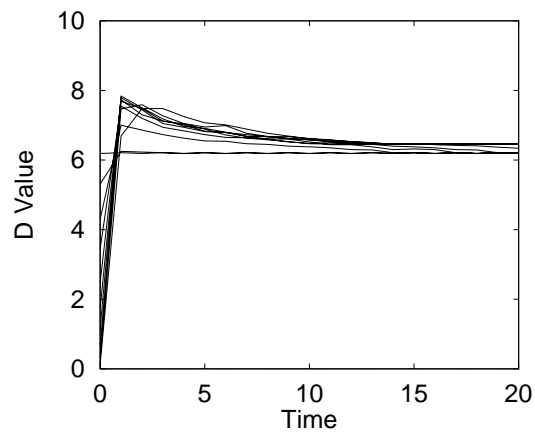
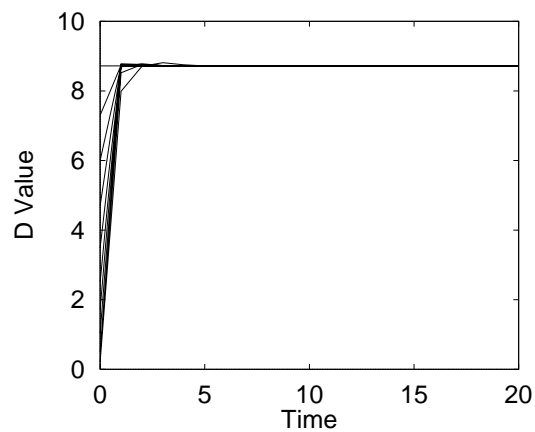
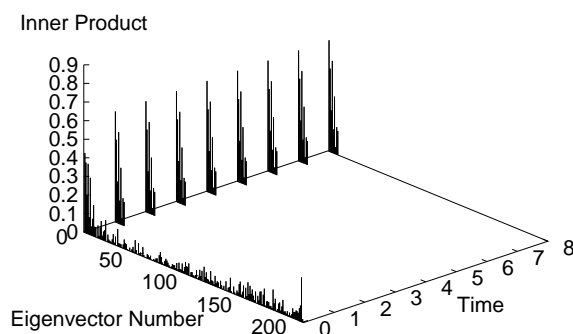
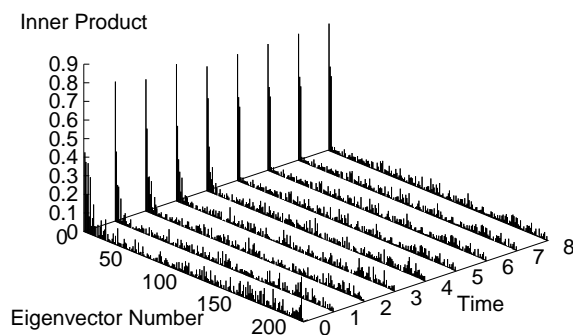
(a) 記憶パターンの想起 ($h = 2.5$)(b) 第二概念パターンの想起 ($h = 2.9$)(c) 第一概念パターンの想起 ($h = 5.0$)

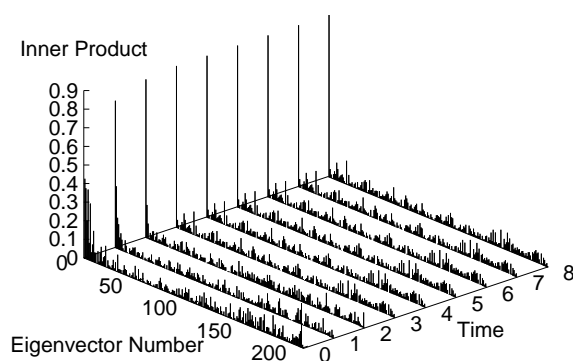
図 5.28: ニューロウインドウ法による選択的想起過程における D 関数の変化 ($N = 1000$, $P_1 = P_2 = P_3 = 3$, $R = 0.49$, $\lambda = 1.0$)。



(a) 記憶パターンの想起 ($h = 2.2$)



(b) 第二概念パターンの想起 ($h = 2.4$)



(c) 第一概念パターンの想起 ($h = 5.0$)

図 5.29: ニューロウインドウ法による選択的想起過程における固有ベクトル成分の変化 ($N = 200, P_1 = 1, P_2 = P_3 = 3, R = 0.49, \lambda = 0.9$)。

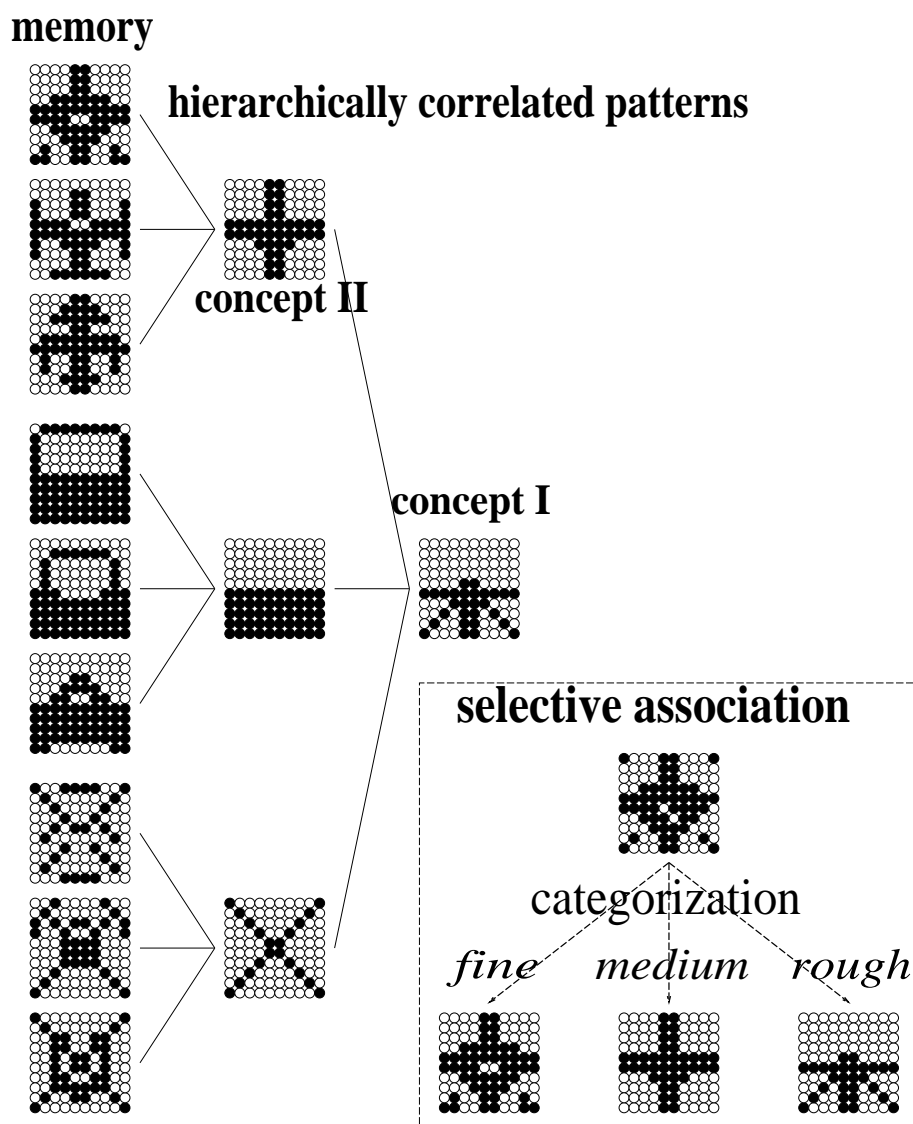


図 5.30: ニューロウインドウ法による階層概念の選択的想起。

5.2.4 ロッシュのベーシックレベルの数理モデル

ニューロウインドウ連想記憶は階層的な概念形成を実現するニューラルネットワークモデルである。この小節では、ニューロウインドウ法による階層概念の形成モデルを、心理学における概念形成に関する知見と対応させて議論を行う。

人間の認識の特徴は似ているものどうしを一つのカテゴリーとして括ることである。そのカテゴリーわけは、ラフなものから細かいものまで臨機応変に使い分けられる。その認識形態に応じて、記憶は図5.31のような階層構造を持つことになる。すなわち、ラフなカテゴリーとしては生物や動物といった分類の器を用意し、細かくは犬、猫などというカテゴリーを使って認識を行う。

この認識の階層性に関して、心理学の立場から、ロッシュによってベーシックレベルという考え方が提唱されている。これは、人間は日常の認識においては、ベーシックレベルと呼ばれる一定の階層レベルでの分類をデフォルトとして用いているという知見である。例えば、人間はスプーンを見たとき、それを「食器」という一段高レベルでの分類をして認識することも可能であるが、通常は「スプーン」という分類レベルで認識する。或は、椅子を見たとき、それを「家具」とラフに分類して認識することも可能であるし、或は「背もたれ付き椅子」という細かい分類で認識することも可能であるが、通常は「椅子」というカテゴリーを用いて認識する。この通常用いられる分類レベルのことをベーシックレベルと呼ぶ。勿論、このベーシックレベルは、個人や文化によっても異なる。例えば、昆虫の専門家ならば、虫を見たとき、その虫の種類まで遡った分類がベーシックレベルになるであろうが、それ以外の人にとっては、特に馴染みのある種以外の虫は全て「虫」というカテゴリーを用いて認識されるであろう。或は、日本語ではベーシックレベルとなる「牛」は、英語では、“cow”, “ox”, “buffalo” など、より細かい分類がベーシックレベルになる。つまり、個人的に、或は文化的により興味や重要度の高い対象に関しては細かい分類がベーシックレベルとなるわけである。

ここで注意しておくが、この議論の枠組を連想記憶モデルと対応する場合、記憶パターンとして想定しているのは、視覚データのようなローデータではなく、特徴抽出後の特徴の集合をパターン化したものである。特徴空間での相関構造を考えると、パターンの相関の中心を概念として定義するのは自然な考え方である。

このベーシックレベルをニューロウインドウ連想記憶モデルを使って解釈すると、次のような説明をすることが可能である。認識において、ある分類レベルがベーシックレベルになることは、ニューロウインドウの閾値 h にデフォルト値が存在することに対応する。ここで、パターンの記憶において、パターンの提示頻度が多いものや、より重要事項に関連するものは記憶が強化される（大きな学習定数を使って記憶される）としよう。この章の議論から、より上位の概念、及び大きな学習定数を使って記憶された概念は線形フローの下流に位置することになる。よって、例えば二つの相関構造を持つクラスターのうち、一つがより強固に記憶されているとすると、ある h のデフォルト値に対して、強固に記憶されたクラスターではより細かい分類のパターンがベーシックレベルとなり、一方、通常のクラスターではラフな分類がベーシックレベルとなる（図5.32）。

具体的に二種類の学習定数を使って二つの階層的相関を持つクラスターを記憶した場

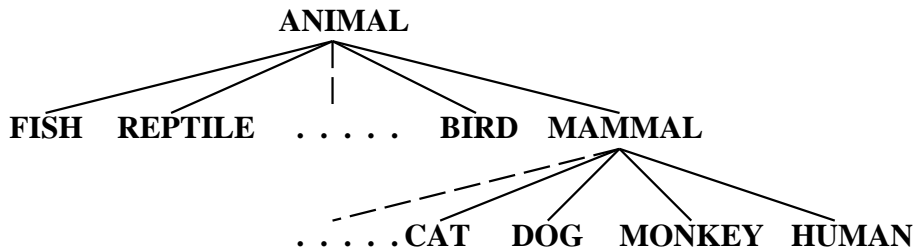


図 5.31: 階層的記憶の例。

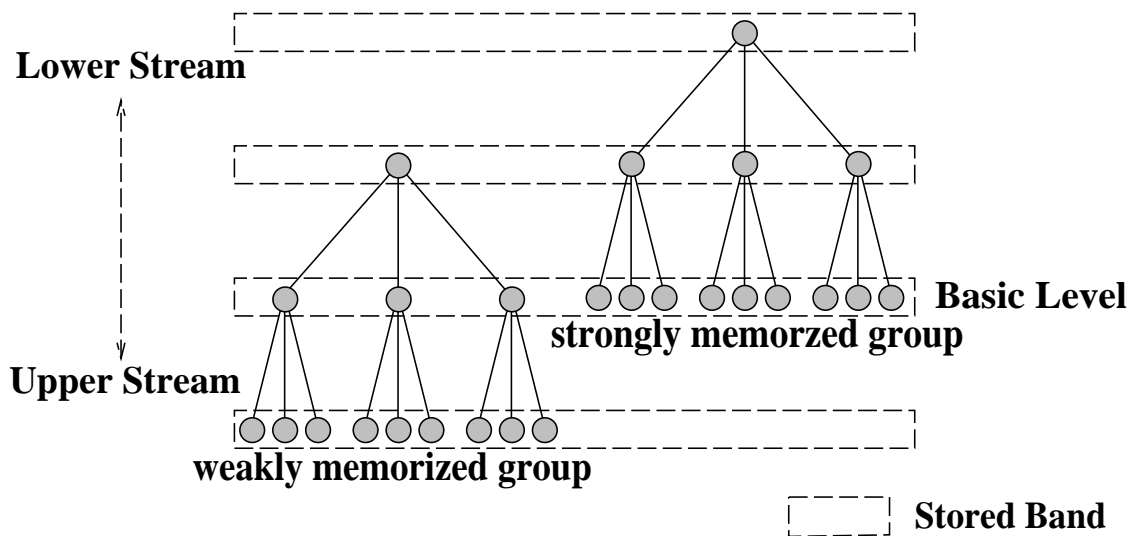


図 5.32: 二つの相関構造を持つクラスターが異なる学習係数で記憶された場合に構成される記憶空間の幾何構造。 h のデフォルト値に対して、強く記憶されたクラスターではより細かい分類のパターンがベーシックレベルとなり、弱く記憶されたクラスターではラフな分類がベーシックレベルとなる。

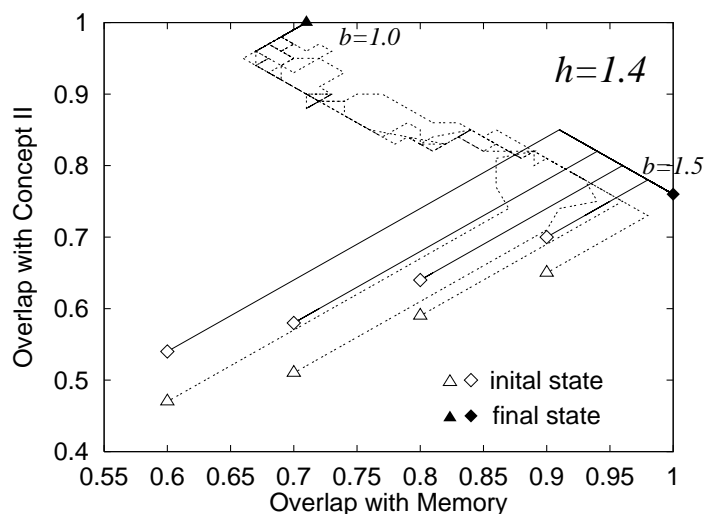


図 5.33: 二つの相関構造 ($R = 0.49, L = 3, P_2 = P_3 = 3, N = 1000$) を持つクラスターが異なる学習係数で記憶された場合の一定のパラメータ h に対する想起過程。強く記憶されたパターンは記憶パターンが想起され、弱く記憶されたパターンは第二概念パターンが想起されている。このように、強く記憶されたクラスターではより細かい分類がなされることになる。

合について、 h を固定した想起を数値実験で行った結果を図 5.33 に示す。このように、同じパラメータ h を用いても、強く記憶されたパターンはより細かいカテゴリーレベルで想起されていることが分かる。

以上のように、ニューロウインドウ法の枠組みでロッシュのベーシックレベルで報告されている現象を再現できることが示された。今後は、ここで示したニューロウインドウに対応する回路網の存在が生理的に確認されることが期待される。

5.3 相互相関連想記憶のニューロウインドウ法

前節においては、自己相関連想記憶を対象にニューロウインドウ法による選択的想起モデルを提案した。この節では、ニューロウインドウ法を相互相関連想記憶に拡張する。前節で述べた通り、ニューロウインドウ法が適用できる例はいくつかあるが、ここでは相関のある系列を記憶した時の記憶系列と概念系列の選択的想起に的を絞って議論を行うことにする。

前章で述べた通り、系列を記憶する連想記憶とは

$$\mathbf{s}^{(p)(0)} \rightarrow \mathbf{s}^{(p)(1)} \rightarrow \dots \mathbf{s}^{(p)(Q-1)} \rightarrow \mathbf{s}^{(p)(0)}$$

という周期 Q の系列 $\{\mathbf{s}^{(p)(\tau)}\}$ を記憶する連想記憶である。相互相関連想記憶では、記憶行列は相互相関行列

$$w_{ij} = \frac{1}{N} \sum_{p=1}^P \sum_{\tau=0}^{Q-1} s_i^{(p)((\tau+1) \bmod Q)} s_j^{(p)(\tau)} \quad (5.21)$$

で与えられる。以下の議論においては、このパターンの系列間に相関がある場合を考える。相関のある系列は

$$\begin{aligned} \mathbf{s}^{(\mathbf{p})(0)} \rightarrow \mathbf{s}^{(\mathbf{p})(1)} \rightarrow \dots \mathbf{s}^{(\mathbf{p})(Q-1)} \rightarrow \mathbf{s}^{(\mathbf{p})(0)} \\ \frac{1}{N} E[\mathbf{s}^{(\mathbf{p})(\tau)} \cdot \mathbf{s}^{(\mathbf{q})(\sigma)}] = \Delta_R(\mathbf{p}, \mathbf{q}) \delta_{\tau\sigma} \end{aligned} \quad (5.22)$$

で定義される。ここで、 $\mathbf{p} = (p_1, \dots, p_L)$ 、 $\mathbf{q} = (q_1, \dots, q_L)$ であり、関数 $\Delta_R(\mathbf{p}, \mathbf{q})$ は

$$\Delta_R(\mathbf{p}, \mathbf{q}) = \begin{cases} 1, & \text{if } p_1 = q_1, \dots, p_L = q_L \\ R^{(L-l)}, & \text{if } p_1 = q_1, \dots, p_l = q_l, p_{l+1} \neq q_{l+1} \\ 0, & \text{if } p_1 \neq q_1 \end{cases} \quad (5.23)$$

で与えられる。また、 $\delta_{\tau\sigma}$ は Kronecker のデルタ関数である。つまり、記憶系列が同じ τ に関して図 5.24 と同じの相関構造を持つ場合を考える。この時、記憶行列は

$$w_{ij} = \frac{1}{N} \sum_{p_1=1}^{P_1} \dots \sum_{p_L=1}^{P_L} \sum_{\tau=0}^{Q-1} s_i^{(\mathbf{p})((\tau+1) \bmod Q)} s_j^{(\mathbf{p})(\tau)} \quad (5.24)$$

で与えられる。ここで N はニューロン数、 P_l は相関の木構造における第 $(l-1)$ 層のノード一つから出る枝の数 (P_1 は木の数) である。このような相関構造をもつ記憶系列に対して、概念系列 $\{s_i^{(p_1, \dots, p_l)(\tau)}\}$ は

$$s_i^{(p_1, \dots, p_l)(\tau)} = \text{sgn} \left(\sum_{p_{l+1}=1}^{P_{l+1}} \dots \sum_{p_L=1}^{P_L} s_i^{(\mathbf{p})(\tau)} \right) \quad (5.25)$$

で定義される。記憶系列と概念系列の例を図 5.34 に示す。

前章で述べた通り、相互相関連想記憶においては、部分反転法を

$$x_i(t+1) = \text{sgn}\left(\sum_{j=1}^N w_{ij}(x_j - \lambda\phi(\tilde{u}_j(t)))\right) \quad (5.26)$$

$$\tilde{u}_j(t) = \sum_{k=1}^N v_{jk}x_k \quad (5.27)$$

$$\phi(\tilde{u}) = \begin{cases} -1, & \text{if } u < -h \\ 0, & \text{if } -h \leq u < h \\ 1, & \text{if } h \leq u \end{cases} \quad (5.28)$$

のように変形したダイナミクスを用いることで、部分反転法と同様に記憶容量を向上させることができる。ここで、行列 $V = [v_{ij}]$ は

$$V = W^T W \quad (5.29)$$

である。また、この V は Q が一定で小さい場合は

$$V = W^Q \quad (5.30)$$

でもよい。よって、この部分反転法のダイナミクスにおいて関数 $\phi(u)$ の閾値 h を制御することで相互相関連想記憶のニューロウインドウ法、すなわち記憶系列と概念系列の選択的想起が実現できると予想される。

数値実験として、二層の相関構造を持つ周期 $Q = 3$ の系列群を記憶したときの記憶系列と概念系列への引き込みの様子を図5.35、図5.36に示す。このように、系列の想起においてもニューロウインドウによる選択的想起ができることが分かる。

前章で、同周期の系列を相関学習で記憶した場合、荷重行列の周期乗はほぼ対称行列になると述べた。この対称行列の固有空間でパターンがどのように表現されるかを示したものが図5.37である。この図から分かるように、ここで得られた固有値の分布、及び記憶／概念系列のパターンと固有ベクトルによる表現は、自己相関連想記憶における固有空間と似ている。すなわち、固有値分布は段差を示しており、概念系列中のパターンは記憶空間の中でも大きい固有値を持つ固有ベクトル群で主に表現されている。このことから、 W^Q (或は $W^T W$) を用いたニューロウインドウ法で選択的想起が実現するメカニズムを自己相関連想記憶とのアナロジーで理解することができる。

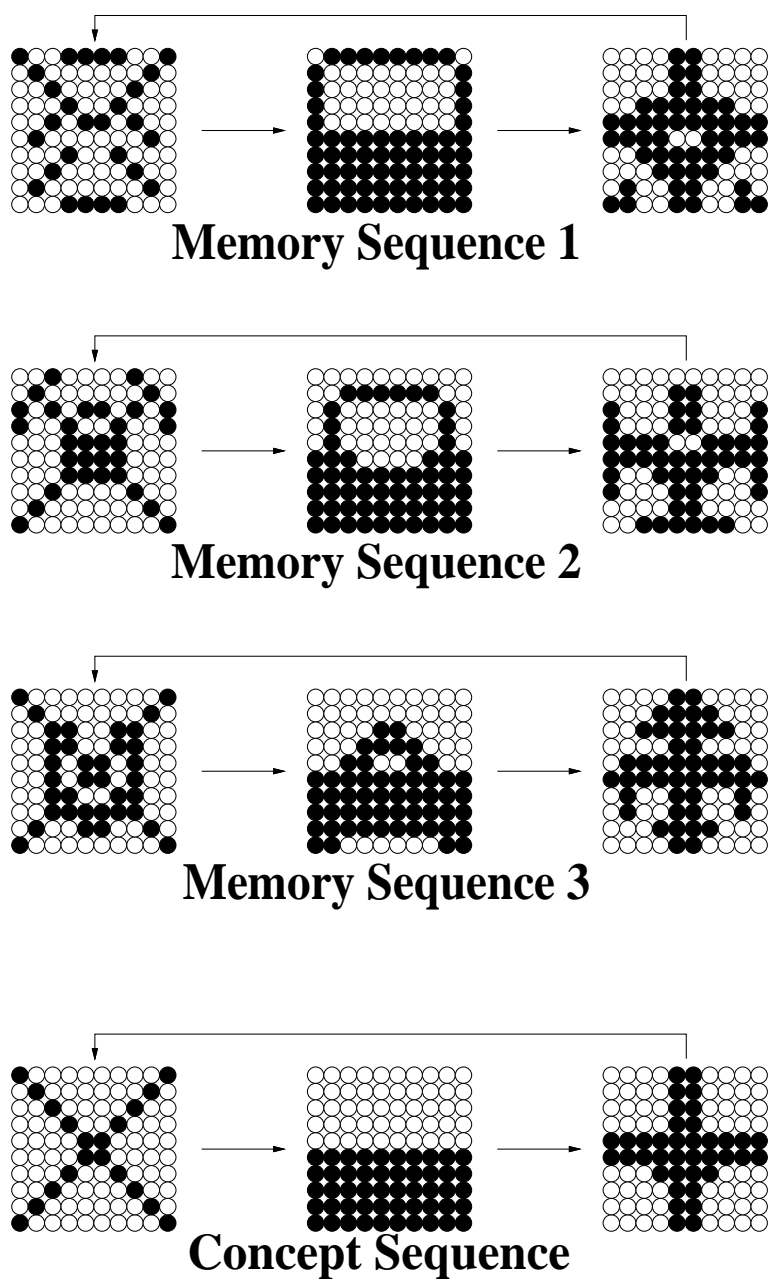


図 5.34: 記憶系列と概念系列の例。

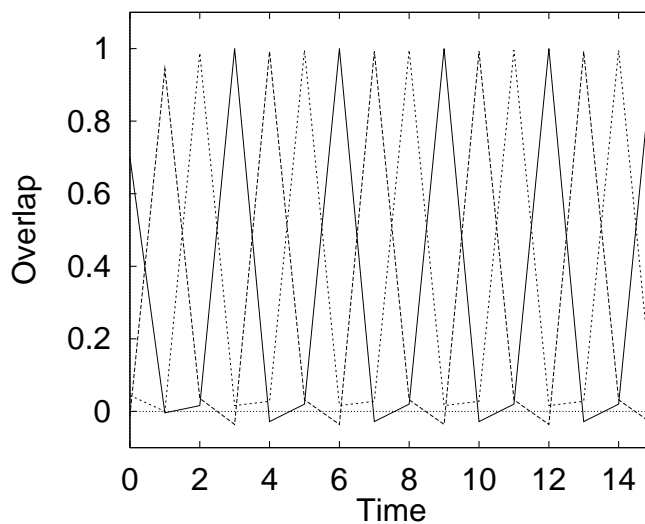
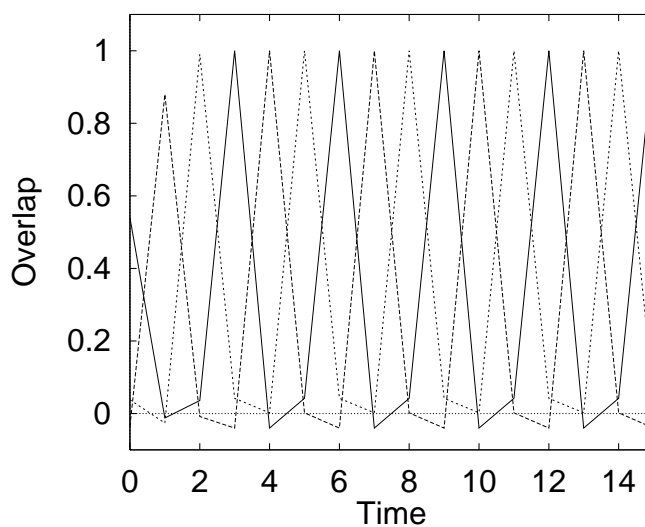
(a) 記憶系列の想起 ($h = 1.0, \lambda = 0.5$)(b) 概念系列の想起 ($h = 10.0, \lambda = 0.5$)

図 5.35: ニューロウインドウ法 (W^Q を使用) を用いた選択的想起過程 ($Q = 3, P_1 = 5, P_2 = 3, R = 0.49, N = 1000$)。 $\tau = 0, 1, 2$ に対応する全パターンとのオーバーラップを表示している。

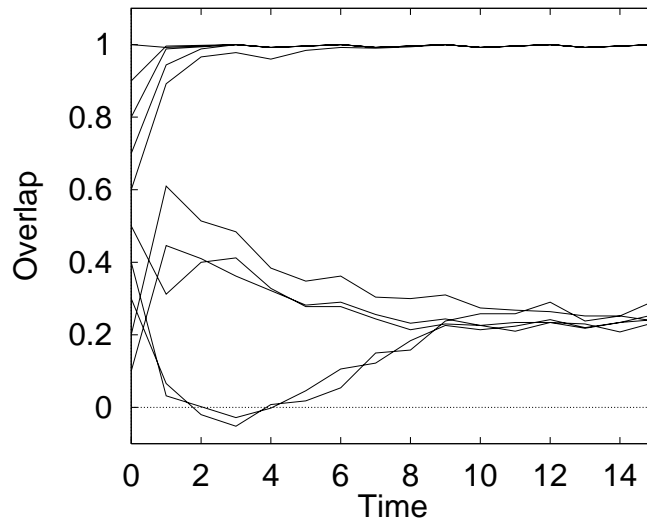
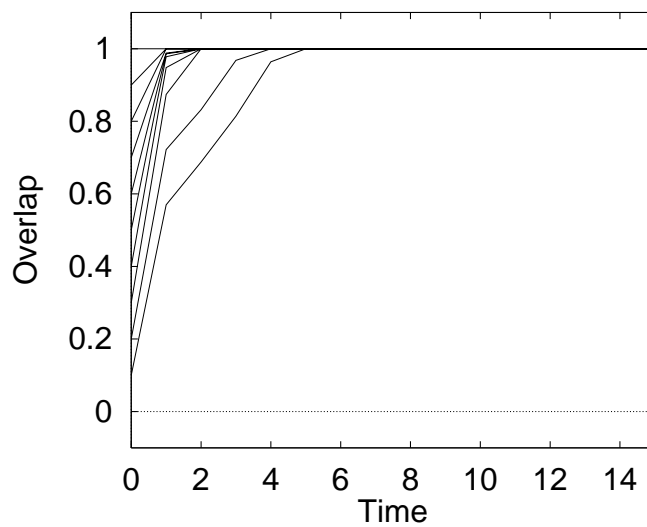
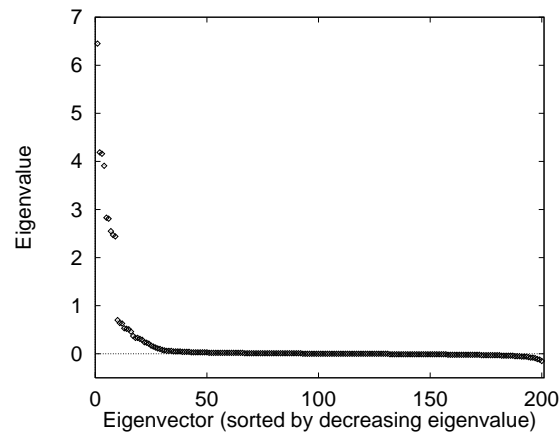
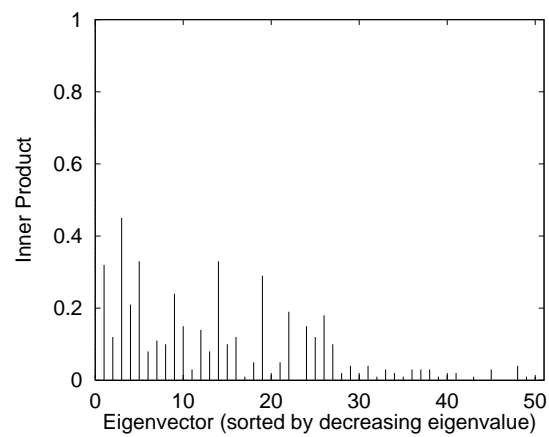
(a) 記憶系列の想起 ($h = 1.0, \lambda = 0.5$)(b) 概念系列の想起 ($h = 10.0, \lambda = 0.5$)

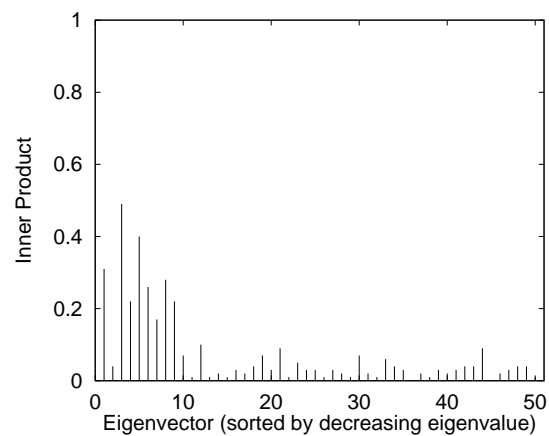
図 5.36: ニューロウインドウ法 (W^Q を使用) を用いた想起における引き込みの様子 ($Q = 3, P_1 = 5, P_2 = 3, R = 0.49, N = 1000$)。異なる初期値に対する軌道とのオーバーラップの経時変化を表示している。



(a) 固有値の分布



(b) 記憶系列のパターンベクトルと固有ベクトルの内積



(c) 概念系列のパターンベクトルと固有ベクトルの内積

図 5.37: 荷重行列を周期乗した行列の対称成分の固有ベクトルと記憶、概念系列のパターンとの内積 ($N = 200, Q = 3, P_1 = 3, P_2 = 3$)。概念系列のパターンは記憶空間中で大きな固有値をもつ固有ベクトル群の成分をより多く持っている。

5.4 (0, 1) 出力系のニューロウインドウ連想記憶

本章で述べてきたように、記憶パターンに相関がある場合、相関のクラスターをなすパターンの平均パターンとして概念パターンを定義することができる。ニューロンの出力が $(-1, 1)$ の二値をとる場合は、部分反転法のパラメータを調節すること（ニューロウインドウ法）により、記憶パターンと概念パターンを選択的に想起できることを示した。この節では、 $(0, 1)$ の二値をとるニューラルネットワークにおいて、ニューロウインドウ法による選択的想起を試みる。

5.4.1 相関構造をもつ $(0, 1)$ パターン

ここでは、簡単のため以下のような相関構造をもつ $(0, 1)$ パターンを記憶する場合を考える。各記憶パターンにおけるニューロンの発火率は aN (N :ニューロン数) に固定する。相関のクラスターが P_1 個あり、それぞれのクラスターは相互には無関係にランダムに生成されるとする。各クラスターは3つの互いに相関のあるパターン $\mathbf{s}^{(p,1)}, \mathbf{s}^{(p,2)}, \mathbf{s}^{(p,3)}$ からなり、その相関は条件

$$\#\{i \mid s_i^{(p,1)} = s_i^{(p,2)} = s_i^{(p,3)} = 1\} = bN \quad (5.31)$$

$$\#\{i \mid s_i^{(p,1)} = s_i^{(p,2)} = 1, s_i^{(p,3)} = 0\} = cN \quad (5.32)$$

$$\#\{i \mid s_i^{(p,2)} = s_i^{(p,3)} = 1, s_i^{(p,1)} = 0\} = cN \quad (5.33)$$

$$\#\{i \mid s_i^{(p,3)} = s_i^{(p,1)} = 1, s_i^{(p,2)} = 0\} = cN \quad (5.34)$$

$$\#\{i \mid s_i^{(p,1)} = 1, s_i^{(p,2)} = s_i^{(p,3)} = 0\} = dN \quad (5.35)$$

$$\#\{i \mid s_i^{(p,2)} = 1, s_i^{(p,3)} = s_i^{(p,1)} = 0\} = dN \quad (5.36)$$

$$\#\{i \mid s_i^{(p,3)} = 1, s_i^{(p,1)} + s_i^{(p,2)} = 0\} = dN \quad (5.37)$$

$$\#\{i \mid s_i^{(p,1)} = s_i^{(p,2)} = s_i^{(p,3)} = 0\} = (1 - b - 3c - 3d)N \quad (5.38)$$

を満たすとする（図 5.38）。ここで $\#$ は集合の要素数を表す。この場合、発火率 a は

$$a = b + 2c + d \quad (5.39)$$

で与えられる。

このようなパターン群においては、 $(-1, 1)$ 出力型ニューラルネットワークの概念パターンに対応する平均パターン（多数決パターン） $\bar{\mathbf{s}}^{(p)} = (\bar{s}_1^{(p)}, \bar{s}_2^{(p)}, \dots, \bar{s}_N^{(p)})^T$

$$\bar{s}_i^{(p)} = \begin{cases} 1, & \text{if } \sum_{q=1}^3 s_i^{(p,q)} \geq 2 \\ 0, & \text{if } \sum_{q=1}^3 s_i^{(p,q)} \leq 1 \end{cases} \quad (5.40)$$

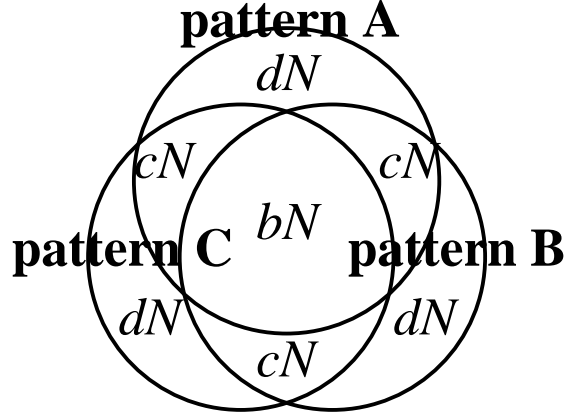


図 5.38: クラスタをなすパターンの相関構造。3つの円は各パターンで興奮するニューロンの集合を表す。

が定義できる。また、その他に、ここでは1出力の持つ特殊性を考慮し、共通パターン $\tilde{\mathbf{s}}^{(p)} = (\tilde{s}_1^{(p)}, \tilde{s}_2^{(p)}, \dots, \tilde{s}_N^{(p)})^T$

$$\tilde{s}_i^{(p)} = \begin{cases} 1, & \text{if } \sum_{q=1}^3 s_i^{(p,q)} = 3 \\ 0, & \text{if } \sum_{q=1}^3 s_i^{(p,q)} \leq 2 \end{cases} \quad (5.41)$$

も考えることにする。

5.4.2 (0, 1) 出力型ニューロウインドウ法

この小節では、前小節で定義したクラスタをなすパターンを相関学習で記憶した場合に、記憶パターン、平均パターン、共通パターンを選択的に想起する方法を提案する。

まず、(0, 1) 出力型のネットワークで連想記憶を構成する一般的な方法を示す。記憶行列は、相関学習、コバリアンス学習の二通りがあるが、ここではコバリアンス学習

$$w_{ij} = \sum_p \sum_q (s_i^{(p,q)} - a)(s_j^{(p,q)} - a) \quad (5.42)$$

を用いる。想起ダイナミクスには、

$$x_i(t+1) = f\left(\sum_{j=1}^N w_{ij}x_j(t) - \theta_i\right) \quad (5.43)$$

$$f(u) = \begin{cases} 1, & \text{if } u > 0 \\ 0, & \text{if } u \leq 0 \end{cases} \quad (5.44)$$

を用いる。閾値の設定法にも複数の方法があるが、ここでは、閾値を $\theta_i = \bar{\theta}(v_i)$ に固定するモデルと、内部状態 u が高い上位 M 個のニューロンを発火させるモデルの二つを考える。(0, 1) モデルでは、単調ダイナミクスでも、閾値固定モデルでは閾値 $\bar{\theta}$ 、発火率固定モデルでは発火ニューロン数 M の設定を変えることで想起の性質に変化を持たせることができる。

パターン間の相関が大きい場合、閾値固定モデルでは、平均パターンは想起しやすく、単調ニューロンの閾値を適切な値にとるだけで想起は可能である(図 5.39)。しかし、閾値 $\bar{\theta}$ を調節しても共通パターンや記憶パターンの想起は実現できない。また、発火率固定モデルでは、 M を小さくすると共通パターンが、 M を大きくすると平均パターンが想起されるので、 M を調節することで共通パターンと平均パターンの選択的想起は可能であるが、記憶パターンの想起はできない。

ここで、単調なダイナミクスでは想起ができないパターンを想起するため、(0, 1) 出力型の部分反転法

$$x_i(t+1) = f\left(u_i(t) - \lambda \sum_{j=1}^N w_{ij} y_j(t) - \theta_i\right) \quad (5.45)$$

$$y_j(t) = f\left(\sum_{k=1}^N w_{jk} x_k(t) - h\right) \quad (5.46)$$

を考える。(−1, 1) 出力型ニューロウインドウ法になぞらえて考えると、 h を調節することで、記憶パターンと平均パターンの選択的想起が可能になると考えられる。 $N = 1000, P_1 = 25, a = 0.2, b = 0.08, c = 0.03, d = 0.06$ の場合について閾値固定モデルで部分反転法を想起に用いた結果を図 5.40 に示す。この図から、パラメータを適切に選ぶと記憶パターンの想起も可能になることが分かる。また、図 5.41 で示す通り、共通パターンについても、 $\lambda < 0$ の部分反転法で想起することができるので、 $\bar{\theta}, h, \lambda$ の調節により記憶パターン、平均パターン、共通パターンの全てを想起することが可能になる。また、発火率固定モデルの場合も、部分反転法において h を適切に選ぶとともに、発火ニューロン数 M を想起したいパターンを構成するニューロン数に合わせると、三種類のパターン全てを選択的に想起することができる(実験結果は省略)。

このような選択的想起が実現されるメカニズムを、閾値固定モデルを例で考えると、次のような説明が成り立つ。図 5.38 のような相関構造をもつパターンをコバリアンス学習で記録した場合、例えば記憶パターン A を入力すると、記憶パターン A では発火せず記憶パターン B、C で発火するニューロン (B C ニューロン) の内部状態が記憶パターン A だけで発火するニューロン (A ニューロン。以下他のニューロンも同様に命名) の内部状態よりも大きくなる。このため、単に閾値を調節するだけでは記憶パターンの想起が不可能になっている。部分反転法を用いると、A B C ニューロンの出力が抑制され、その結果 B C ニューロンと A ニューロンの内部状態の大きさの順序を逆転させることができるので、閾値の調節により記憶パターンの想起が可能になる。一方、共通パターンを想起する場合は、閾値を上昇させ A B C ニューロンのみを発火させようとする、他のニューロンからの入力が無くなるので、その結果 A B C ニューロンの内部状態も低下し、全ニューロンが非発火状態となる。そこで、部分反転法で $\lambda < 0$ とし、内部状態が

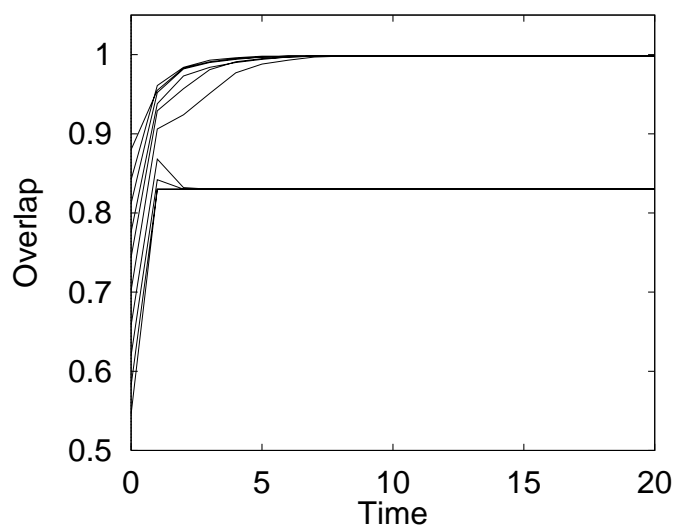


図 5.39: 単調な想起過程における平均パターン（概念パターン）との一致度の変化 ($N = 1000$, $P_1 = 25$, $a = 0.2$, $b = 0.8$, $c = 0.3$, $d = 0.6$, $\bar{\theta} = 0.1$)。

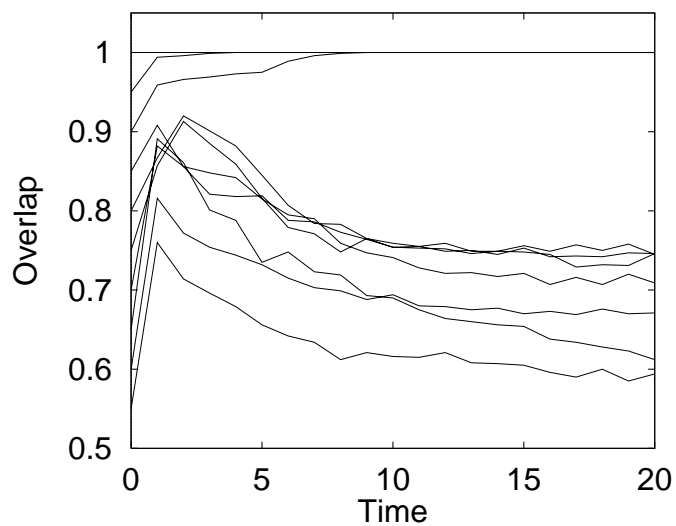


図 5.40: 部分反転法を用いた想起過程における記憶パターンとの一致度の変化 ($N = 1000$, $P_1 = 25$, $a = 0.2$, $b = 0.8$, $c = 0.3$, $d = 0.6$, $\bar{\theta} = 0.03$, $\lambda = 0.6$, $h = 0.09$)。

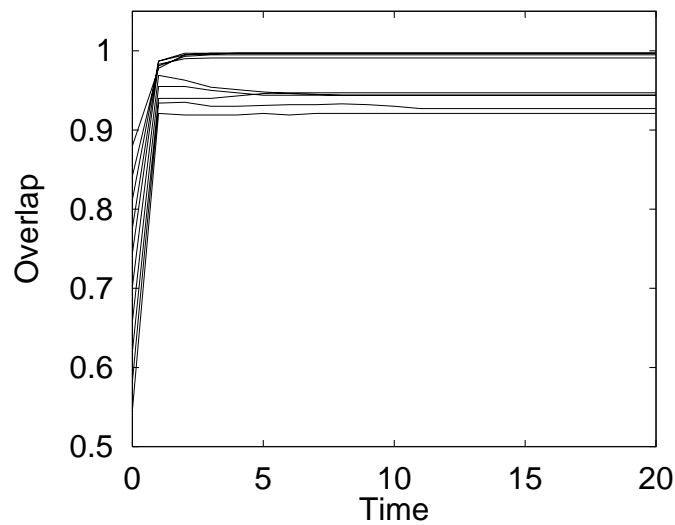


図 5.41: 部分反転法を用いた想起過程における共通パターンとの一致度の変化 ($N = 1000$, $P_1 = 25$, $a = 0.2$, $b = 0.8$, $c = 0.3$, $d = 0.6$, $\bar{\theta} = 0.03$, $\lambda = 0.6$, $h = 0.09$)。

高いニューロンの出力を増強すれば、ABCニューロンのみの発火状態を安定化することができるわけである。ここで説明した内容は、第6章で紹介する有限系解析を用いると、定量的に議論することが可能である。有限系解析の実際の計算例は、第7章の連続時間ダイナミクスの議論において掲載している。

5.5 オートマトンにおける冗長度低減モデル

本章のこれまでの議論では、ニューロウインドウ法をホップフィールド型のネットワークについて導入した。ホップフィールドネットワークでは外部入力に初期状態としてのみ与えられるので、その力学系は自励的である。しかし、一般の神経回路網による情報処理を議論する上では、外部信号を受け取るニューロン層を考慮したシステムを考える必要がある。本節においては、そのような外部入力層を含む層状のネットワークの例としてニューロオートマトンを取り上げ、ニューロウインドウ法の原理を用いてどのような情報処理が行えるかを検討する。

5.5.1 問題設定

まず、この節で提案するモデルの背景にある具体的な問題意識を述べる。

入力層と内部層をもつニューラルネットワークを用意し、その二層間に相互想起型連想記憶を埋め込めば、図 5.42 のように if-then ルールを記憶させることができる。例えば、「もし内部状態が B の時に A という入力を得られたら内部状態を C にする」というルールを覚えさせたければ、入力層から内部層への結合に B から C への相互相関、内部層内の結合に A から C への相互相関を記憶すれば良い。

互いに関連のある if-then ルールが記憶されていくと、連想の系列が生成されていく。例えば、「もし内部状態が B の時に A という入力を得られたら内部状態を C にする」というルールと「もし内部状態 C の時に D という入力を得られたら内部状態を E にする」というルールが記憶されると、図 5.42 のようなルールの結合がなされる。

多くの if-then ルールを覚えたシステムは、環境から入力を受けながら、ルールに従って内部状態を次々変えていく。そのような環境下における経験の中で、最初に学習したルールの中に冗長なものがあることが見つかるケースが考えられる。例えば、図 5.42 の例でいくと、内部状態が C の時、ほとんどの全てのケースで D が入力されたとすると、内部状態 A の時に B が入力された段階で C をとばして E を連想する方が効率的である。実際、人間の情報処理においてはこのような効率化がよく見られる。しかしながら、こういった処理は従来の連想記憶の記憶法と想起法の組み合わせでは実現できない。

この節では、そのような冗長なステップをとばすモードと、最初に覚えたルール通りに想起するモードの切り替えのできるシステムを、本章で導入したニューロウインドウ法の考え方を利用して実現する方法を紹介する。

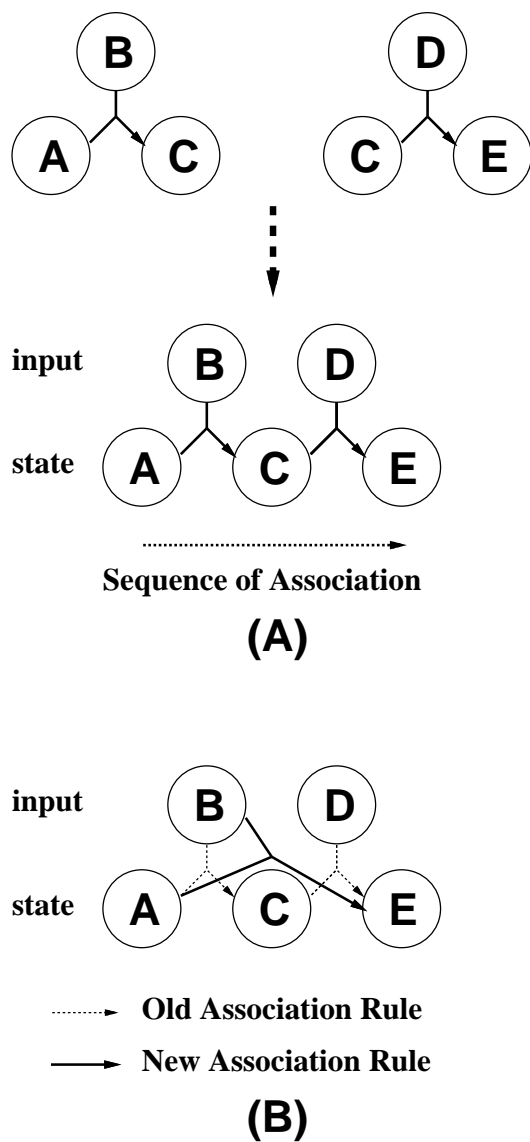


図 5.42: 連想系列の生成: (A) ルールの結合; (B) 冗長なステップの省略。

5.5.2 ネットワークの構造と初期学習

今、入力層と内部層はそれぞれ N 個のニューロンから構成されており、外部からの入力信号、内部状態は、それぞれ N 次元のベクトル \mathbf{x}, \mathbf{y} で表現されるとする。ここでは、情報の表現法としてスパースコーディング [5] を採用する。すなわち、ベクトル \mathbf{x}, \mathbf{y} を、 $M = aN (a \ll 1)$ 個の (+1) 成分と $(N - M)$ 個の 0 成分からなるベクトルに制限する。これは、各層におけるニューロンの発火率が一定に保たれることを意味する。

今、ニューラルネットワーク型連想記憶に、オートマトンを学習させたいとする。学習させるオートマトンは図 5.43 のように、一つの初期状態と、いくつかの終状態がある。そして、複数の内部状態を周回するようなループはないとする。以下に、このようなオートマトンをニューラルネットワークで実現するためにはどのようなシステムを組めば良いかについて述べる。

ネットワークがオートマトンとして働くためには状態遷移が入力と内部状態の両方に依存しなくてはならない。これを実現する方法はいくつかあるが、ここでは、図 5.44 のように、入力信号と内部状態の積を表現する中間層を用意し、そこからの入力

$$\bar{x}_{ij} = x_i y_j \quad (5.47)$$

によって内部状態が遷移するようにネットワークを組む方法を採用することにする。このネットワークに、図 5.43 のようなオートマトンを記憶させるには、中間層の第 jk ニューロンから内部層の第 i ニューロンへの結合を

$$v_{ijk} = \frac{c_1}{M^2} \sum_{(\nu, \mu) \in A} (s_i^\mu - a)(r_j^{(\nu, \mu)} - a)(s_k^\nu - a) \quad (5.48)$$

内部層の第 j ニューロンから第 i ニューロンへの結合を

$$w_{ij} = \frac{c_2}{M^2} \sum_{\nu} (s_i^\nu - a)(s_j^\nu - a) \quad (5.49)$$

とし、想起ダイナミクスを

$$y_i(t+1) = 1_L \left(\sum_{j=1}^N w_{ij} y_j(t) + \sum_{j=1}^N \sum_{k=1}^N v_{ijk} \bar{x}_{jk}(t) \right) \quad (5.50)$$

$$1_L(u) = \begin{cases} 1, & \text{if } u \geq L \\ 0, & \text{if } u < L \end{cases} \quad (5.51)$$

とすればよい。ここで、 $(\nu, \mu) \in A$ は、学習するオートマトンが状態 s^ν から状態 s^μ へ直接のパスをもつことを示し、その状態遷移を促す入力を $r^{(\nu, \mu)}$ と記述している。また、内部ニューロン間の結合 w_{ij} は、記憶したパターンを安定化する結合である。 c_1, c_2 は学習定数で $c_1 > c_2$ である。今、各パターンは M 個のニューロンの興奮パターンとして表現されているので、 L は興奮する細胞数を M に固定するように適応的に変化する閾値であるとする。それによって、パターンが一つだけ想起されることになる。

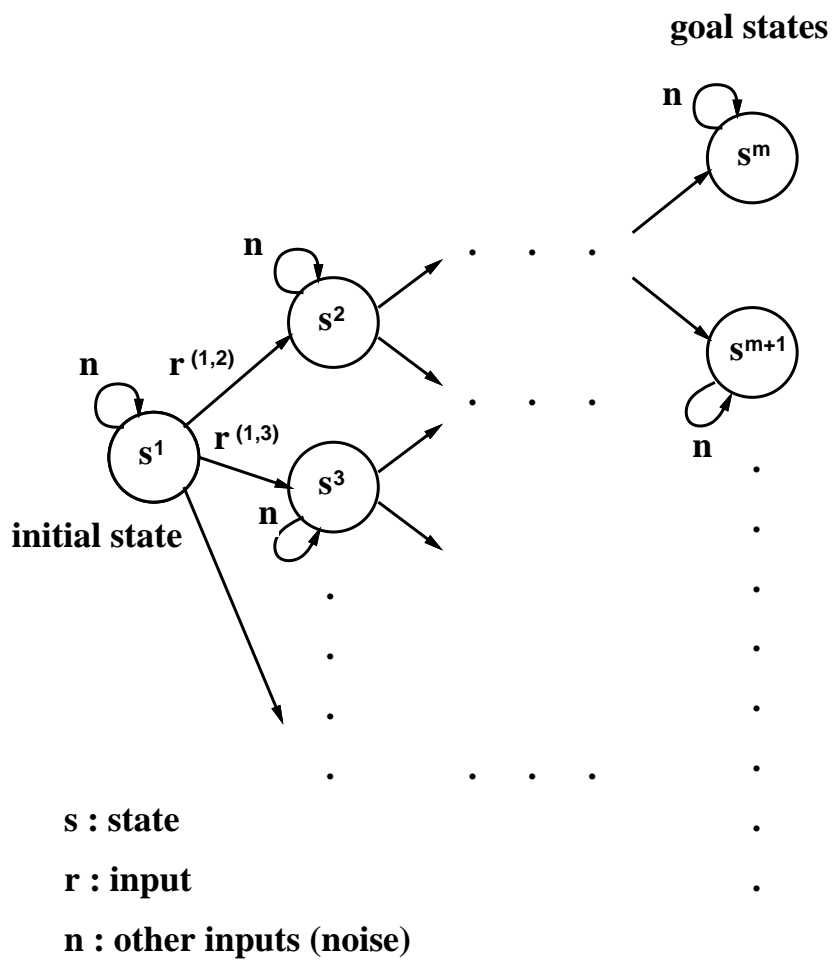


図 5.43: 本節の議論でニューラルネットワークが学習するオートマトン。初期状態からいくつかの分岐を経て終状態へと至る。複数の状態を周回するようなループはないとする。

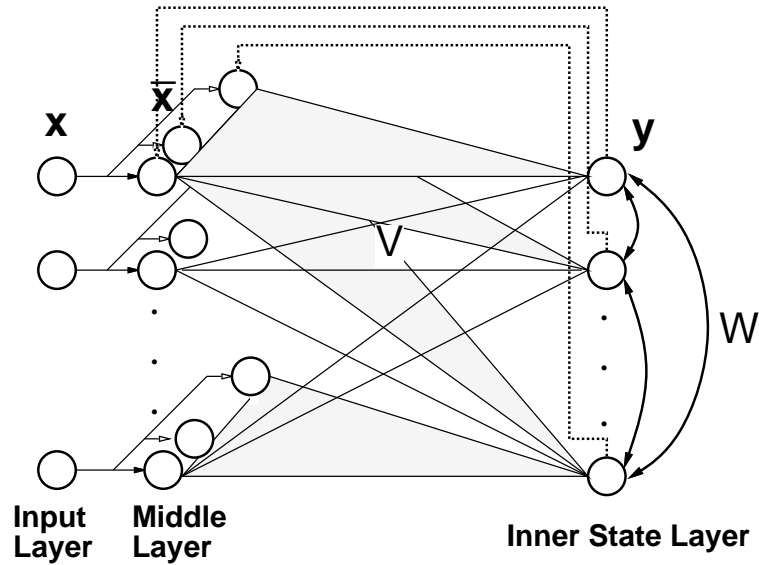


図 5.44: オートマトンを学習するネットワークの構造。入力と内部状態のオンオフ状態を表現する $N \times N$ 個の中間層ニューロンと内部層ニューロン間に連想のための相関構造を埋め込む。

5.5.3 経験からの学習

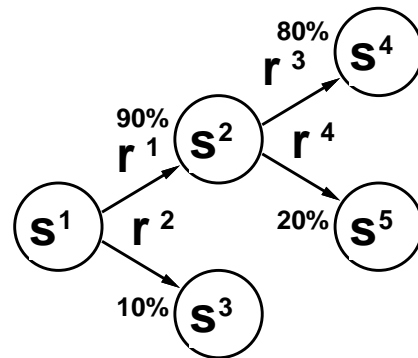
どの状態が冗長であるかは、環境、すなわち入力信号の確率的性質によって決まるものである。よって、冗長な状態をとばして「効率的」な連想を行うには、経験によって環境の性質を学習していかななくてはならない。この考えに基づいて、次のような学習則を提案する。

一試行間に、システムは初期状態から出発し、外部入力を受けとりながら、最終状態に達するまで、内部状態を変えていくとする。そして、各試行の状態遷移において、ネットワークは、相関学習則 (Hebb 則) に従って、結合荷重を変化させていくようにする。

ここで、学習則としては、内部層の第 k ニューロンが興奮している時に入力層の第 j ニューロンが興奮した場合、その後で内部層の i 番目のニューロンが興奮する条件つき確率に荷重 v_{ijk} が比例するように学習する。これは、中間層の jk 番目ニューロンが興奮した後で内部層の第 i ニューロンが興奮すれば結合 v_{ijk} を強め、興奮しなければ結合を弱めることで実現される。この時、結合荷重は、

$$v_{ijk} = \frac{c_1}{M^2} \sum_{(\nu, \mu') \in B} P(\mu' | \nu, \mu) (s_i^{\mu'} - a) (r_j^{(\nu, \mu)} - a) (s_k^\nu - a) \quad (5.52)$$

となる。ここで、 $(\nu, \mu') \in B$ はオートマトンにおいて状態 s^ν の下流に状態 $s^{\mu'}$ があることを示す。また、 $P(\mu' | \nu, \mu)$ は、状態 s^ν で入力 $r^{(\nu, \mu)}$ が入力されたという条件のもとで、



Strength of Correlation

$\mathbf{s}^1, \mathbf{r}^1$ to \mathbf{s}^2 : 1.0	$\mathbf{s}^1, \mathbf{r}^2$ to \mathbf{s}^3 : 1.0
$\mathbf{s}^1, \mathbf{r}^1$ to \mathbf{s}^4 : 0.8	$\mathbf{s}^2, \mathbf{r}^3$ to \mathbf{s}^4 : 1.0
$\mathbf{s}^1, \mathbf{r}^1$ to \mathbf{s}^5 : 0.2	$\mathbf{s}^2, \mathbf{r}^4$ to \mathbf{s}^5 : 1.0

図 5.45: 条件付き確率学習の簡単な例。

状態 \mathbf{s}^i がその後に経由される条件つき確率を表す。このような学習則を条件付き確率学習 (CP 学習: Conditional Probability Learning) と呼ぶことにする。この学習によるパターン間相関の記憶強度の構造を、図 5.45 に簡単な例を用いて示しておく。

5.5.4 想起過程

単調増加の閾値素子を用いた想起ダイナミクスでは、経験による学習をした後も、得られる想起の結果は同じである。ここで我々が目標としているのは、いつも同様の入力を得るような冗長な状態は省略して想起することである (図 5.46)。以下の議論では、ニューロウインドウ法を応用することで、経験からの学習結果を埋め込んだネットワークが系列を効率的に想起できるようになることを示す。

ある入力から状態を連想する時、より達成される確率の高い状態を表すニューロンほど大きな入力を受け、より達成される確率が低い状態を表すニューロンはより小さな入力を受けることになる。であるから、ニューロウインドウ法のメカニズムを利用し、大きい内部状態を持つニューロンを不活性化することができれば、ある確率 P 以上で達成される状態を表現するニューロンの活性を抑えることが可能となる。これによって、確率 P 以上で達成される状態を除いた中で最も実現確率の高い状態を想起するということが可能である。具体的には、ニューロウインドウの閾値 R (前節までの議論における h に対応するパラメータ) を

$$R = (1 - a)^3 c_1 P \quad (5.53)$$

とし、想起ダイナミクス

$$y_i(t+1) = g(u_i^1(t), u_i^2(t)) \quad (5.54)$$

$$u_i^1(t) = \sum_{j=1}^N w_{ij} y_j(t) + \sum_{j=1}^N \sum_{k=1}^N v_{ijk} \bar{x}_{jk}(t) \quad (5.55)$$

$$u_i^2(t) = \sum_{j=1}^N v_{ijk} \bar{x}_{jk}(t) \quad (5.56)$$

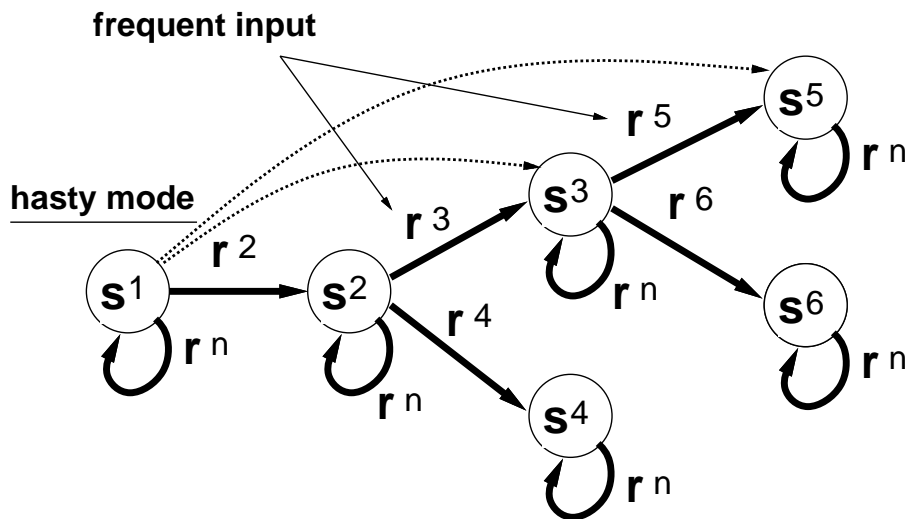
$$g(u^1, u^2) = \begin{cases} 0, & \text{if } u^2 \geq R \\ 1, & \text{if } u^1 \geq L, u^2 < R \\ 0, & \text{if } u^1 < L, u^2 < R \end{cases} \quad (5.57)$$

を使用すればよいことになる。ここで、 L は興奮するニューロン数を M にするような閾値である。つまり、確率 P 以上で実現されるパターンは、そのパターンを表すニューロンの内部状態の平均が $(1 - a)^3 c_1 P$ 以上になるので、そのようなニューロンを不安定化するわけである。但し、この R の設定は、スパースコーディングによる内部状態の分布の偏りや、記憶間の相関などより多少揺らぐことになる。

また、 $c_1 > c_2$ であることを考慮すると、上のダイナミクスを簡略化し、想起ダイナミクス

$$y_i(t+1) = f\left(\sum_{j=1}^N w_{ij} y_j(t) + \sum_{j=1}^N \sum_{k=1}^N v_{ijk} \bar{x}_{jk}(t)\right) \quad (5.58)$$

$$f(u) = \begin{cases} 0, & \text{if } u \geq R \\ 1, & \text{if } R > u \geq L \\ 0, & \text{if } u < L \end{cases} \quad (5.59)$$



s^i : inner state of the network

r^i : input to the network

(r^n : noise input)

Memorized Rules

$(s^1 r^2) \rightarrow s^2$ $(s^2 r^3) \rightarrow s^3$ $(s^2 r^4) \rightarrow s^4$

$(s^3 r^5) \rightarrow s^5$ $(s^3 r^6) \rightarrow s^6$

図 5.46: 冗長な状態をスキップする連想。いつも同じ入力を受ける状態は省略する。

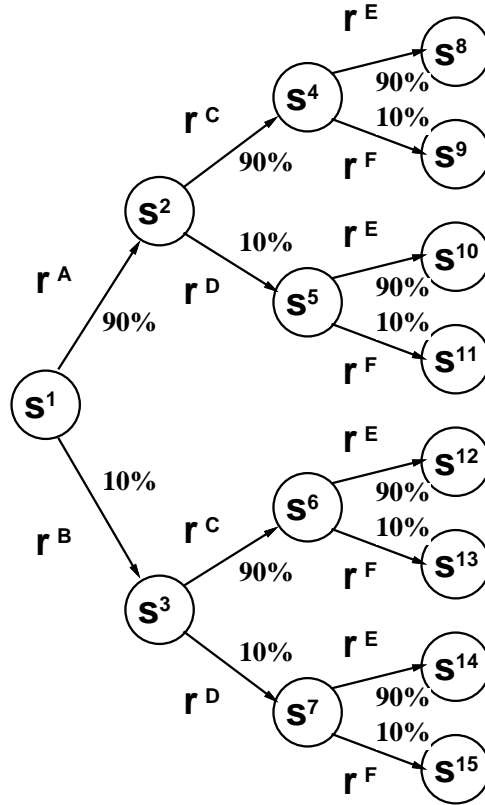
を用いても同じ結果が得られると期待される。

上のダイナミクスは確率 P の大きい状態を不活性化しているが、確率 P が 1 に近い状態遷移は上で議論した冗長な状態遷移に対応し、このルールに従った状態遷移はここで定義している状態遷移の効率化になっている。但し、状態遷移の効率化のもう一つの定義として、ある確率 P 以上で達成される状態の中で最も遠い (P の小さい) 状態に行くという定式化も考えられる。この定義に基づいた状態遷移の効率化を実現するには、閾値 L を先に $(1-a)^3 c_1 P$ より小さめに設定し、興奮するニューロン数を M 個に制約するパラメータとして閾値 R を使えばよい。

5.5.5 実験

数値実験として、システムに図 5.47 に示したオートマトンを、上で述べてきた学習則を用いて学習させた。学習の結果得られたネットワークを用い、パラメータ R を変えて想起を行った実験結果を図 5.47 に示す。実験では学習係数を $c_1 = 3$, $c_2 = 1$ とおいた。また、ここでは、内部層の初期状態を s^1 とし、入力信号は 1 単位時間だけ入力して、内部ダイナミクスの収束後に得られた内部層の状態を実験結果として示している。更に、 L は発火ニューロンを $M (= 10)$ に固定する可変閾値である。

実験結果から分かるに、 R を小さくすると、ほとんどいつも同じ入力を受けるような途中の状態をとばし、より遠くの状態まで一度に遷移するようになる。これは、ここで定義している状態遷移の効率化が実現されていることを意味する。また、ノイズの入った入力をシステムに加えると、状態スキップをあまりしなくなるという結果も得られた。これは、内部ニューロンへの入力が小さくなることに起因している。このように、入力の不確かな時は状態スキップをせず、より「慎重」になるという生物のシステムとしては好ましい性質が自然な形で導入されている。



(a) 数値実験で学習対象として用いたオートマトンの構造

Memorized Internal State s^1 :

```

-----
-----+-----
-+-+-----+-----+
+-----+-----+-----+
    
```

Memorized Internal State s^2 :

```

-----+-----+-----+-----
-----+-----+-----+-----
-+-+-----+-----+-----
-----+-----+-----+-----
    
```

Memorized Internal State s^4 :

```

-----+-----+-----+-----
-----+-----+-----+-----
-----+-----+-----+-----
-----+-----+-----+-----
    
```

Memorized Internal State s^8 :

```

-----+-----+-----
-----+-----+-----
-+-+-----+-----+-----
-----+-----+-----+-----
    
```

(b) 記憶パターン ($M = 10, N = 100$)

Association from Input r^A (0% noise)

R=4.0

```

-----+-----+-----+
-----+-----+-----+
-++-----+-----+-----
-----+-----+-----+

```

R=3.0

```

-----+-----+-----+
-----+-----+-----+
-----+-----+-----+
-----+-----+-----+

```

R=1.6

```

-----+-----+-----+
-----+-----+-----+
-+-----+-----+-----
-----+-----+-----+

```

(c) R を変えた時の想起パターン

Association from Input r^A (20% noise)

R=3.0

```

-----+-----+-----+
-----+-----+-----+
-++-----+-----+-----
-----+-----+-----+

```

(d) ノイズを含む入力に対する想起パターン

Association from Input r^C (0% noise)

R=4.0

```

-----+-----+-----+
-----+-----+-----+
-+-----+-----+-----
+-----+-----+-----+

```

(e) 記憶に関係のない入力を与えられた時の想起パターン

図 5.47: 数値実験の結果: (a) 学習するオートマトン; (b) 記憶パターン; (c) R が小さい場合は途中の状態をとばしてより先の状態が想起される; (d) ノイズを含む入力を与えられると、状態遷移の跳躍は小さくなる; (e) 記憶に関係ない入力に対しては状態は変化しない。

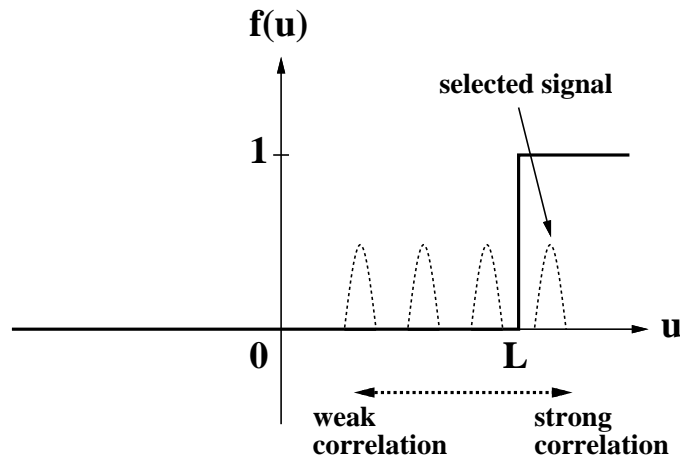


図 5.48: 一般の想起ダイナミクス。

5.5.6 考察

今まで提案されたほとんどの神経回路モデルにおいては、個々のニューロンは他のニューロンからの入力をうけとり、その荷重和が閾値以上であるかどうかで、興奮するか否かを決定する形式をとっている。今、 j 番目のニューロンから i 番目のニューロンへの結合荷重を w_{ij} とすると、個々のニューロンの数理的な表現は

$$x_i(t+1) = 1_L \left(\sum_j w_{ij} x_j(t) \right) \quad (5.60)$$

となる。しかし、Hebb 型の学習で相関情報がネットワークに埋め込まれている場合、式 (5.51) で表されるような活性化関数では、相関の強い情報しか取り出すことができない。つまり、ネットワーク自体は相関の強度の順序も含めた情報が載っているのに、それを十分生かすことが出来なくなっている。これを生かすには、想起過程に工夫を設ける必要がある。その工夫の一つが、ニューロンの活性化関数を閾値型 (図 5.48) から窓型 (図 5.49) にすることで、ある範囲の相関の情報だけを取り出すというニューロウインドウ法のアプローチであると言える。

本章で述べてきたニューロウインドウ法は、欲しい相関のピークをパラメータコントロールによって取り出すというものであった。相関情報を有効に使うもう一つのアプローチとして考えられるのが、相関のピークを動かすという方法である。例えば、マクロな状態区分がいくつかあり、そのマクロな状態によって、システムの状態遷移ルールを変更したい場合を考える。最も単純には、そのマクロな状態だけ連想記憶システムを用意するという方法もあるが、マクロな状態の種類が非常に多い場合、システムをそれだけ別に用意するのは、負担が大きくなり効率が悪い。そこで、図 5.50 のような結合によるネットワークを考える。このモデルでは、ニューロン間の結合に分岐を設けている。そ

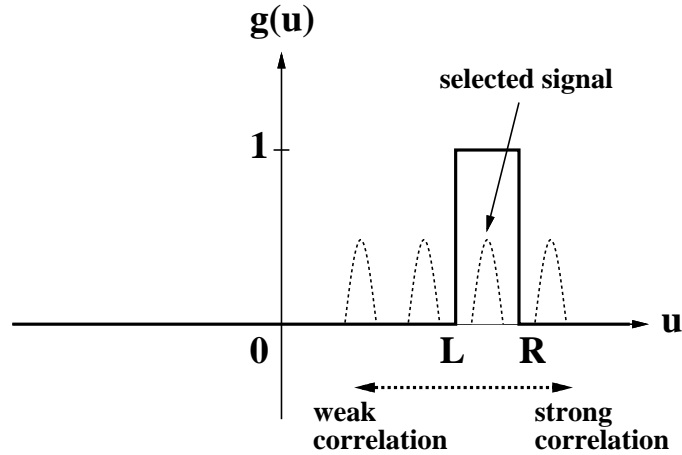


図 5.49: 一定の相関をもつパターンを選択的に想起するダイナミクス。

して、それぞれの枝が、上層のニューロンによってオンオフ制御を受けている。つまり、上層の k 番目のニューロンの状態を $x_k^{(1)}$ 、下層の i 番目のニューロンの状態を $x_i^{(2)}$ 、上層の k 番目のニューロンによる制御を受ける下層の j 番目のニューロンから i 番目のニューロンへの結合を w_{ij}^k としたとき、この結合をもつネットワークのダイナミクスを

$$x_i^{(2)}(t+1) = f\left(\sum_k \sum_j x_k^{(1)} w_{ij}^k x_j^{(2)}(t)\right) \quad (5.61)$$

で与える。その結果、マクロな情報を担う上層のニューロンが、ミクロな情報を担う下層のニューロン間の情報処理を制御することができる。

ここで、例えば、入力がマクロな情報 $r^{(\text{mac})}$ とミクロな情報 $r^{(\text{mic})}$ からなる時、マクロな状況が $r^{(\text{mac})\lambda}$ で、ミクロな情報として入力 $r^{(\text{mic})\mu}$ が与えられた時は s^ν を連想しなさいというルールをネットワークに覚えさせたければ、図 5.51 のような構造をもつネットワークを用い、結合荷重を

$$w_{ij}^k = \sum_{\lambda, \mu, \nu} r_k^{(\text{mac})\lambda} s_i^\nu r_j^{(\text{mic})\mu} \quad (5.62)$$

とすればよい。このダイナミクスでは、マクロな状態によって、相関のピークを動かし、下層での連想性がコントロールされる (図 5.52)。この結合を用いれば、マクロな状況によって、状態遷移則を変えるオートマトンが生成できる。さらには、図 5.53 のような構造を持ったネットワークを使えば、オートマトンを階層的に重ねていくこともできる。

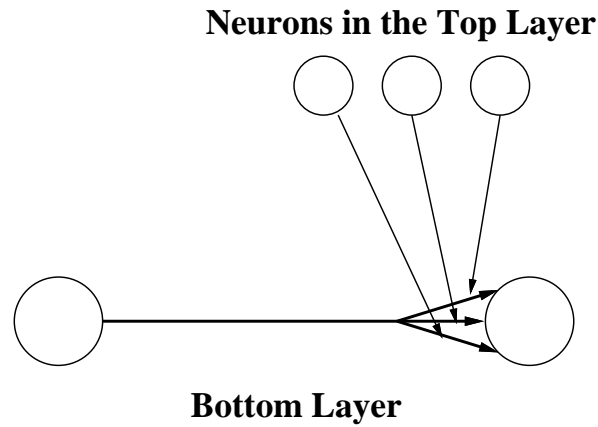


図 5.50: 上層ニューロンによる下層ニューロンの情報伝達の制御。下層のニューロン間の結合に枝分かれがあり、その枝を通じた信号伝達のオンオフを上層のニューロンが制御する。

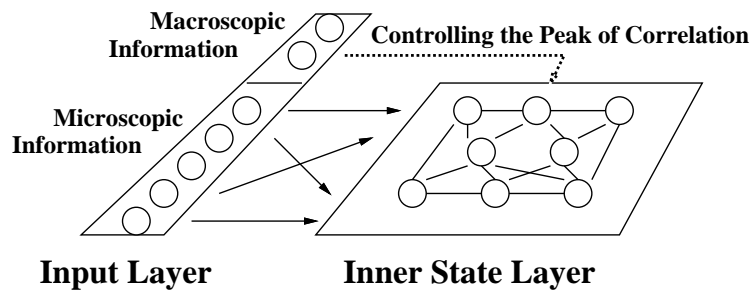


図 5.51: マクロな情報によって連想ルールを変えるネットワーク。マクロな状態を表現するニューロンを上層ニューロンとし、連想の制御を行わせる。

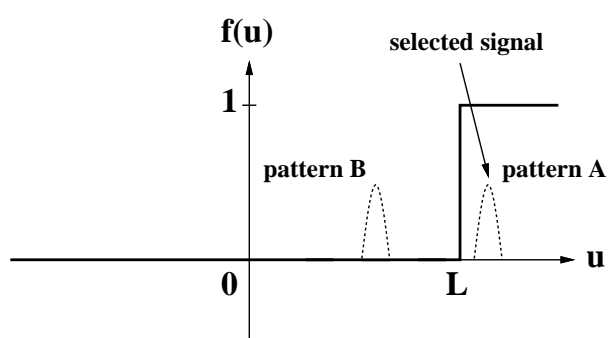
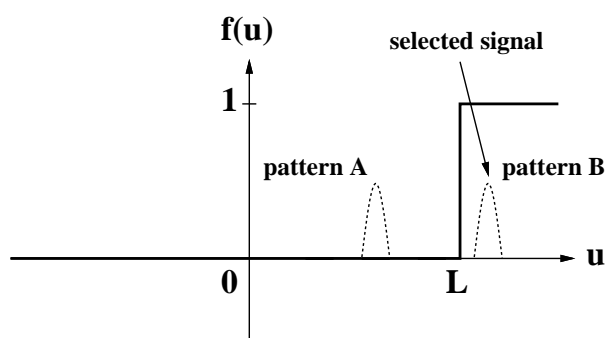
Under Macroscopic State I**Under Macroscopic State II**

図 5.52: マクロな状態によるネットワークの連想性の制御。マクロな状態の違いによって、下層ニューロンの相関のピークが交換され、異なった連想を行うことができる。

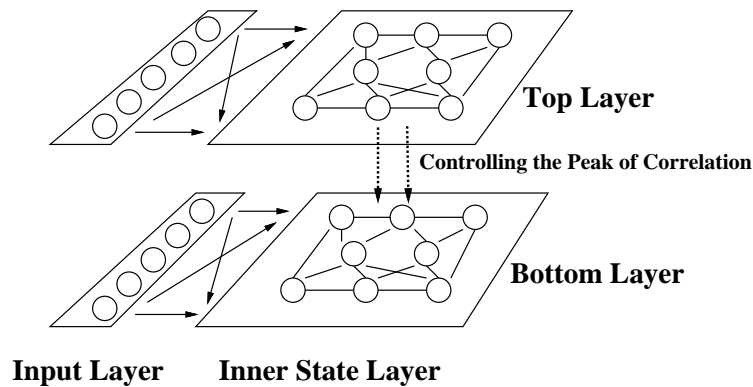


図 5.53: 階層的な情報処理を行うニューロオートマトン。マクロな状態に過去の履歴によるオートマトンの構造を埋め込むと、階層的なオートマトンシステムを構成することができる。

5.6 まとめ

本章では、連想記憶における選択的想起法としてニューロウィンドウ法を提案した。ニューロウィンドウ法を用いることで、学習係数の異なるグループの選択的想起、階層的相関構造を持つパターン群に対する記憶パターンと概念パターンの選択的想起等が実現される。また、系列を記憶する相互相関連想記憶でも記憶軌道と平均軌道の選択的安定化が実現できる。更に、ニューロウィンドウ法はニューロオートマトンにおける状態遷移の効率化にも有効に機能することを示した。

この章で議論した階層的な知識を埋め込んだ階層的情報処理は、脳のシステムに良く見られる現象であるから、そのような情報処理を担う構造として、この章で述べたようなシステムの原理が実際の脳でも使われている可能性は十分あると考えられる。

最後に、この章で述べた研究の今後の課題として、次の三つを挙げておく。まず、ここで提案したシステムをさらに拡張することで、より高度な情報処理を行うようなシステムを模索することができると考えられる。次に、ここであげたシステムの工学的な分野での実用性を検討してみるのも面白いであろう。最後に、生理学的な立場から、ここで挙げたようなシステムと実際の脳のシステムとの対応づけをしていくことも、脳の高次機能の研究への足掛かりとして、一つの切り口に成ると考える。

第 6 章：連想記憶の統計理論解析

前章において、階層的相関をもつパターンを記憶した相関型連想記憶において、ニューロウインドウ法を用いて記憶パターンと概念パターンを選択的想起するモデルを幾何学的視点から提案した。本研究では、このモデルの性質を S/N 解析の視点から検討する。

自己相関連想記憶 [8, 17, 22, 31] に関しては、今までに数多くの理論解析が行われている。その中で、安定点解析には、レプリカ法 [6, 7] や SCSNA [40] などの解析法が提案されており、一方、ダイナミクスについては統計神経力学 [4] などの解析が行われている。

本章においては、安定点解析に SCSNA を、力学解析に統計神経力学を用い、ニューロウインドウ連想記憶の統計的理論を導く。まず、第 1 節において、S/N 解析の基礎となる有限系の解析について述べる。次に、第 2 節では、ニューロウインドウ連想記憶の SCSNA を求める。第 3 節においては、ニューロウインドウ連想記憶の統計神経力学を求めるための基本として部分反転法の統計神経力学を導く。最後に、第 4 節において、ニューロウインドウ連想記憶の統計神経力学を求める。

6.1 ニューロウインドウ連想記憶の有限系解析

ニューロン数 N が無限大の極限であっても、記憶パターン数が有限の場合には、想起ダイナミクスを有限個の微分方程式に落とすことができる [26, 27, 39]。この節では、その有限系の解析手法について述べる。

今、記憶パターン数が M 個であるとする。この時、各ニューロンについて、1 番目のパターンから M 番目のパターンまでで発火するかしないかを記述していくことができる。例えば、 $M = 3$ の場合を考えると、あるニューロンは 1 番目のパターンでは発火、2 番目のパターンでは非発火、3 番目のパターンでは発火するならば、そのニューロンに $(+-+)$ というラベルを貼ることができる。このラベルを全ニューロンに貼ることになると、 N 個のニューロンは、そのラベルの違いによって、 2^M 種類に分類することができる。

2^M に分類されたニューロンのうち、同じ分類に属するニューロンは、全パターンで同じ発火パターンを示すので、自己相関学習を行った結果得られる荷重は全く同一になる。であるから、同じ分類に属するニューロンは、想起ダイナミクスにおいて常に同じ発火パターンを示すことになる。よって、各分類 α ($\alpha = 1, 2, \dots, 2^M$) の出力状態 x_α によって系の状態遷移は表せることになる。

今、グループ α のニューロン数を n_α とする。また、記憶パターン \mathbf{s}^p の各分類 α での出力値を s_α^p とする。更に、現状態 $\mathbf{x}(t)$ とパターン \mathbf{s}^p の相関 m を

$$m_p(t) = \frac{1}{N} \sum_{p=1}^M \mathbf{s}^p T \mathbf{x}(t) \quad (6.1)$$

とする。自己相関連想記憶の荷重行列 W におけるグループ α とグループ β 間の結合荷重は

$$w_{\alpha\beta} = \frac{1}{N} \sum_{p=1}^M s_{\alpha}^p s_{\beta}^p \quad (6.2)$$

であるから、ニューロンの内部状態 $u_{\alpha}(t)$ は

$$u_{\alpha}(t) = \sum_{\beta} w_{\alpha\beta} n_{\beta} x_{\beta}(t) \quad (6.3)$$

$$= \sum_{\beta} \left(\frac{1}{N} \sum_{p=1}^M s_{\alpha}^p s_{\beta}^p n_{\beta} x_{\beta}(t) \right) \quad (6.4)$$

$$= \frac{1}{N} \sum_{p=1}^M s_{\alpha}^p \sum_{\beta} n_{\beta} m_{\beta} \quad (6.5)$$

で与えられる。

よって、想起ダイナミクスは

$$x_{\alpha}(t+1) = \text{sgn} \left(\sum_{p=1}^M s_{\alpha}^p m_p(t) \right) \quad (6.6)$$

$$m_p(t) = \frac{1}{N} \left(\sum_{\alpha=1}^{2^M} n_{\alpha} s_{\alpha}^p x_{\alpha}(t) \right) \quad (6.7)$$

で与えられることになる。

また、部分反転法の場合は、

$$x_{\alpha}(t+1) = \text{sgn} \left(\sum_{p=1}^M s_{\alpha}^p (m_p(t) - \tilde{m}_p(t)) \right) \quad (6.8)$$

$$m_p(t) = \frac{1}{N} \left(\sum_{\alpha=1}^{2^M} n_{\alpha} s_{\alpha}^p x_{\alpha}(t) \right) \quad (6.9)$$

$$y_{\alpha}(t) = \phi \left(\sum_{p=1}^M s_{\alpha}^p m_p(t) \right) \quad (6.10)$$

$$\tilde{m}_p(t) = \frac{1}{N} \left(\sum_{\alpha=1}^{2^M} n_{\alpha} s_{\alpha}^p y_{\alpha}(t) \right) \quad (6.11)$$

となる。この有限系の式が、統計解析のシグナル部分に対応し、以下の節で扱う議論の基礎となる。

6.2 ニューロウインドウ連想記憶の SCSNA

SCSNAとはself-consistent signal-to-noise analysisの略である [40]。この方法は、ニューロンの内部状態のうち、想起するパターン由来の項 (シグナル項) とそれ以外の項 (ノイズ項) を分割し、ノイズ項に対して正規分布の仮定をおくことで、系の安定点の定量的評価を行うものである。この節では、前章で紹介したニューロウインドウ法の SCSNA を求め、得られた理論値と実験値の比較を行って得られた SCSNA の妥当性を検証する。

6.2.1 SCSNA

ニューロウインドウ法は、前章で紹介したとおり様々な情報処理に応用可能であるが、ここではその中でも興味深い概念パターンと記憶パターンの選択的想起モデルを取り扱う。但し、多層の相関構造を持つ場合は解析が複雑になるので、ここでは最も単純な相関の階層構造が二層の場合について限定して議論を行う。

階層構造を二層に限定することによる簡略化と、他の統計解析に関する論文との整合性のため、本章の以下の議論では前章からノテーションを以下のように多少変更することにする。

記憶パターンを ξ^μ で表し、相関のある集合を A_1, A_2, \dots, A_p で表すことにする。各集合 A_p は M 個の成分を持ち、記憶パターン間の相関は

$$\frac{1}{N} E[\xi^\mu \cdot \xi^\nu] = \begin{cases} 1, & \text{if } \mu = \nu \\ r^2, & \text{if } \mu \neq \nu, \xi^\mu \in A_p, \xi^\nu \in A_p \\ 0, & \text{if } \xi^\mu \in A_p, \xi^\nu \in A_q, p \neq q \end{cases} \quad (6.12)$$

を満たすとする。この時、概念パターン $\bar{\xi}^p$ は

$$\bar{\xi}_i^p = \text{sgn}\left(\sum_{\xi^\mu \in A_p} \xi_i^\mu\right) \quad (6.13)$$

で定義される。また、荷重行列は

$$w_{ij} = \frac{1}{N} \sum_{\mu=1}^{aMN} \xi_i^\mu \xi_j^\mu - aNM \delta_{ij} \quad (6.14)$$

となる。ニューロウインドウ法の想起ダイナミクスは

$$x_i(t+1) = \text{sgn}\left[\sum_{j=1}^N w_{ij} \left(x_j(t) - \phi\left(\sum_{k=1}^N w_{jk} x_k(t)\right)\right)\right] \quad (6.15)$$

$$\phi(u) = \begin{cases} -\lambda, & \text{if } u < -h \\ 0, & \text{if } -h \leq u < h \\ \lambda, & \text{if } h \leq u \end{cases} \quad (6.16)$$

の閾値 h を制御することになる。

このように定式化されたニューロウインドウ連想記憶の SCSNA は

$$m_\mu^{(\tau)} = \int_{-\infty}^{\infty} Dz \langle \xi^\mu Y_\tau(z) \rangle_{\xi^\mu} \quad (\tau = 1, 2) \quad (6.17)$$

$$q_{\tau\rho} = \int_{-\infty}^{\infty} Dz \langle Y_\tau(z) Y_\rho(z) \rangle_{\xi^\mu} \quad (\tau, \rho = 1, 2) \quad (6.18)$$

$$U_\tau = \frac{1}{\sigma_\tau} \int_{-\infty}^{\infty} Dz z \langle Y_\tau(z) \rangle_{\xi^\mu} \quad (\tau = 1, 2) \quad (6.19)$$

$$Y_1(z) = \phi \left(\sum_{A_p} \xi^\nu m_\nu^{(2)} + \Gamma_{11} Y_1(z) + (\Gamma_{12} - aM) Y_2(z) + \sigma_1 z \right) \quad (6.20)$$

$$Y_2(z) = \text{sgn} \left(\sum_{A_p} \xi^\nu (m_\nu^{(2)} - m_\nu^{(1)}) + (\Gamma_{21} + aM) Y_1(z) \right. \\ \left. + (\Gamma_{22} - aM) Y_2(z) + \sigma_2 z \right) \quad (6.21)$$

$$H = \begin{pmatrix} 0 & A \\ B & C \end{pmatrix} \quad (6.22)$$

$$A_{ij} = U_1 \delta_{ij} + r^2 U_1 (1 - \delta_{ij}) \quad (6.23)$$

$$B_{ij} = -U_2 \delta_{ij} - r^2 U_2 (1 - \delta_{ij}) \quad (6.24)$$

$$C_{ij} = U_2 \delta_{ij} + r^2 U_2 (1 - \delta_{ij}) \quad (6.25)$$

$$K = \begin{pmatrix} K^{(1,1)} & K^{(1,2)} \\ K^{(2,1)} & K^{(2,2)} \end{pmatrix} = (I - H)^{-1} \quad (6.26)$$

$$V = \begin{pmatrix} V^{(1,1)} & V^{(1,2)} \\ V^{(2,1)} & V^{(2,2)} \end{pmatrix} \\ = \begin{pmatrix} K^{(2,1)} & K^{(2,2)} \\ K^{(2,1)} - K^{(1,1)} & K^{(2,2)} - K^{(1,2)} \end{pmatrix} \quad (6.27)$$

$$\Gamma_{\mu\nu} = a(\text{Tr} V^{(\mu,\nu)} + r^2 \sum_{i=1}^M \sum_{j \neq i}^M V_{ij}^{(\mu,\nu)}) \quad (6.28)$$

$$\sigma_\tau^2 = \sum_{\mu=1}^2 \sum_{\nu=1}^2 \left[a q_{\mu\nu} \left(\sum_{i=1}^M \sum_{j=1}^M V_{ij}^{(\tau,\mu)} V_{ij}^{(\tau,\nu)} \right) + r^4 \sum_{i=1}^M \sum_{j=1}^M \sum_{k \neq i}^M \sum_{l \neq j}^M V_{ij}^{(\tau,\mu)} V_{kl}^{(\tau,\nu)} \right. \\ \left. + 2r^2 \sum_{i=1}^M \sum_{j=i}^M \sum_{k=1}^M V_{ik}^{(\tau,\mu)} V_{jk}^{(\tau,\nu)} \right] \quad (6.29)$$

$$Dz = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-z^2}{2}\right) dz \quad (6.30)$$

と計算される (付録 B)。ここで、 $\xi^\mu \in A_p$ であり、 $A, B, C, K^{(\mu,\nu)}, V^{(\mu,\nu)}$ は $M \times M$ 行列である。

6.2.2 理論値と実験結果の比較

この小節においては、上で求めた SCSNA の平衡点を数値計算で求め、シミュレーションの結果と比較する。ニューロウインドウ法においては、記憶率 a , 相関 r , パラメータ h, λ 等、多くの要素が絡むので、本論文ではそれぞれのパラメータをふりながら、SCSNA がどの場合もシミュレーションを説明できているかどうかを確認することにする。

前述したように、ニューロウインドウ法の計算機シミュレーションでは、部分反転法のパラメータ h を制御することで、記憶パターンと概念パターンを選択的に想起できることが確認されている。その様子が SCSNA でも見られるかどうかを調べるため、 $r = 0.7, M = 3$ の場合について、記憶率を増やしていった時の系の安定点の変化を示す(図 6.1)。図 6.1 のように、部分反転法 ($h = 1.6, \lambda = 0.8$) を用いたときの平衡点と記憶パターンの相関は、記憶率 $aM = 0.135$ 付近まではほぼ 1 である。しかし、それ以上のパターンを記憶させようとすると、相関は急激に 0.7 に落ち込む。以下の議論で記憶容量を論じる時は、この m が 1 から急激に遠ざかる点を記憶容量として定義する。一方、通常ダイナミクス ($h = \infty$) で概念パターンを想起する場合は、平衡点と概念パターンの相関が 1 で、記憶率を上昇させるとその相関は 1 から徐々に遠ざかる。つまり、記憶率がある程度のレベルまでは、 h の値を操作することで平衡点を記憶パターンと概念パターンに選択的にスイッチできることを示している。

選択的想起の様子を更に明示するため、 h を変化させた時の記憶パターンと概念パターンの記憶容量を調べた結果を図 6.2 に示す。このように、シミュレーションで得られたとおり、 h を小さくとると記憶パターンが安定化され、 h を大きくすると概念パターンが安定化される様子が SCSNA でも再現されている。

また、異なる相関 r における記憶パターンの容量を図 6.3 に示す。図 6.3 から分かるように、記憶パターンは r が小さい方が記憶容量が大きい、逆に概念パターンは r の小さい方が記憶容量が大きいというシミュレーションの結果が、得られた SCSNA でも忠実に再現されている。

最後に、 λ の大きさを変えた時の系の変化を見る。図 6.4 から分かる通り、ほぼ全ての r に渡って λ が小さい時の方が大きい記憶容量が実現できる。この結果はシミュレーションと SCSNA で一致している。よって、ニューロウインドウ法においては小さい λ の方が好ましいと言える。しかし、 r が 0.9 前後まで大きい時には小さい λ では選択的想起が完全に不能になるが、大きい λ を用いると記憶容量は非常に小さいが理論上選択的想起が実現される。この事実は $N = 1000$ 程度のシミュレーションではあまりに小さい記憶容量を検出することはできなかったため、SCSNA によって新たに得られた結果である。

以上のように、得られた SCSNA の理論値はニューロウインドウ連想記憶のシミュレーションの結果をよく説明していることが確認された。

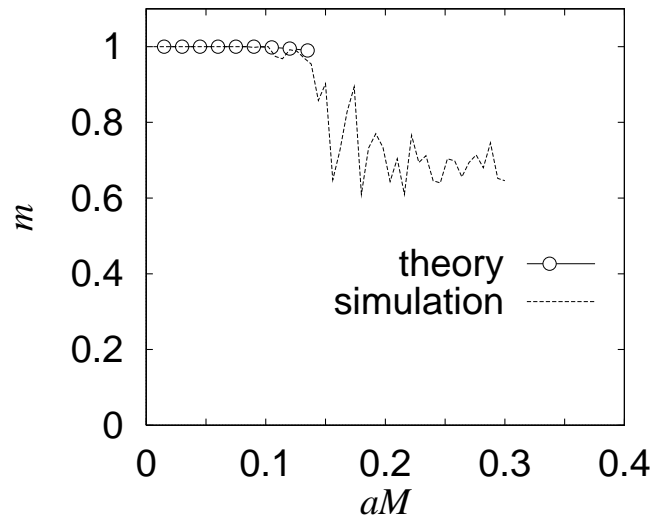
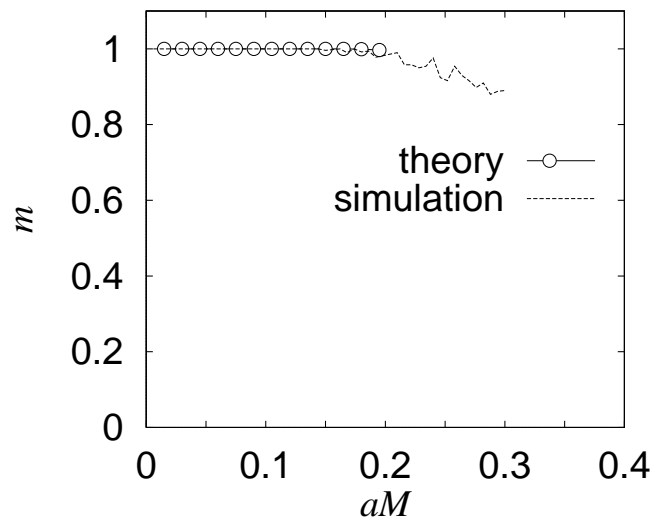
記憶パターンの安定性 ($h = 1.5$)概念パターンの安定性 ($h = 2.6$)

図 6.1: 記憶パターンと概念パターンの安定性 ($r = 0.7, \lambda = 0.8, M = 3$)。

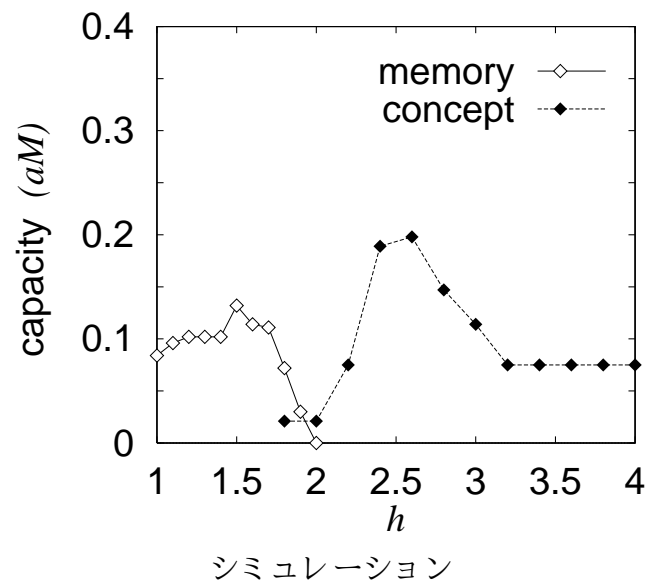
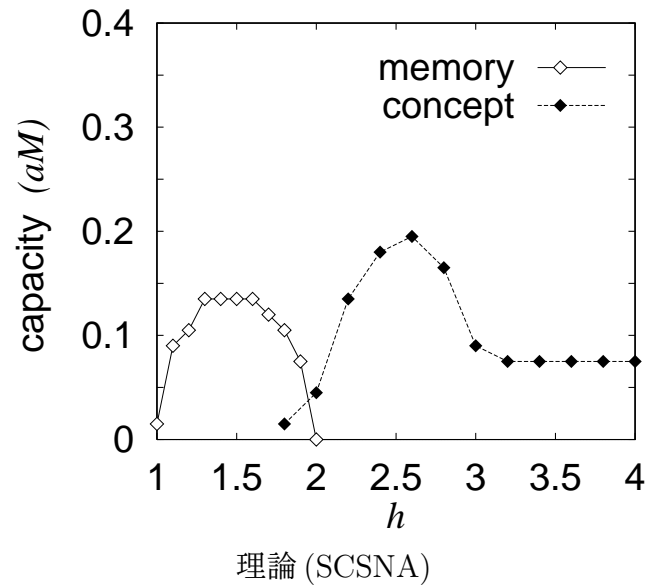


図 6.2: ニューロウインドウ法で h を変化させた時の記憶パターンと概念パターンの安定性 (記憶容量) ($\lambda = 0.8, M = 3$)。

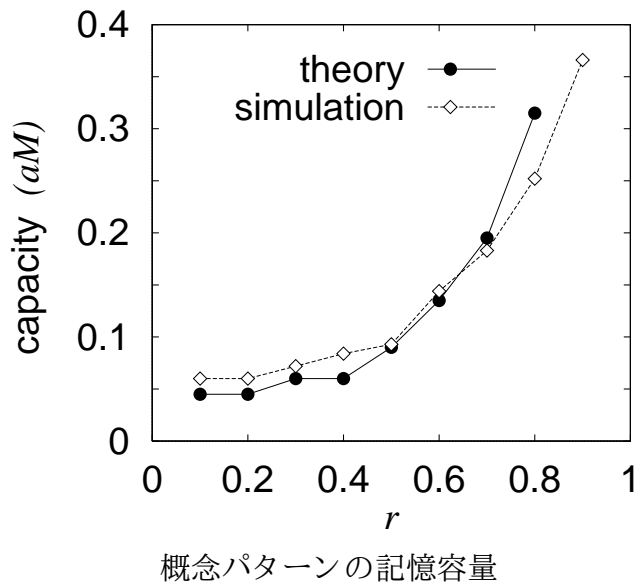
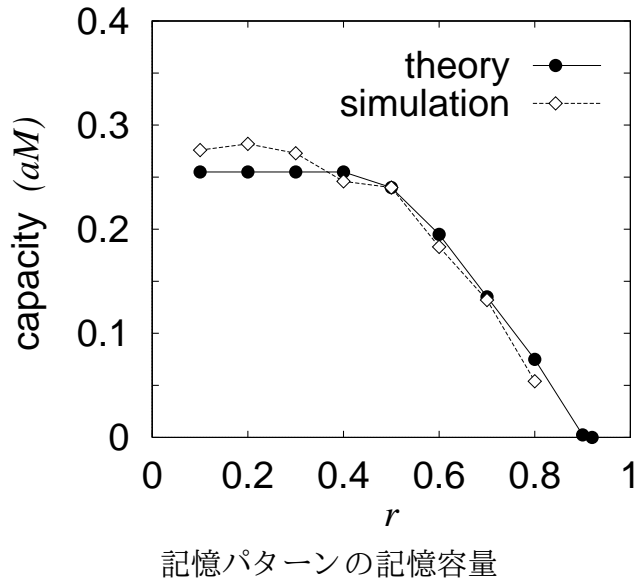


図 6.3: パターンの相関の違いによる記憶/概念パターンの記憶容量の変化 ($\lambda = 0.8, M = 3$)。パラメータ h は各相関での最適値を選んでいる。

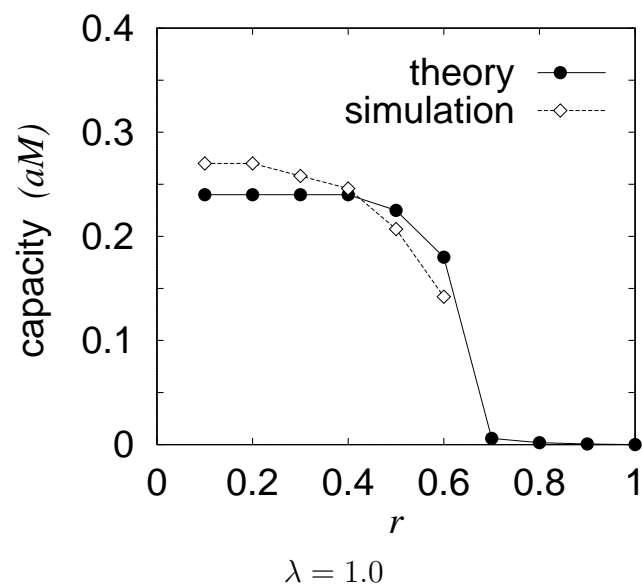
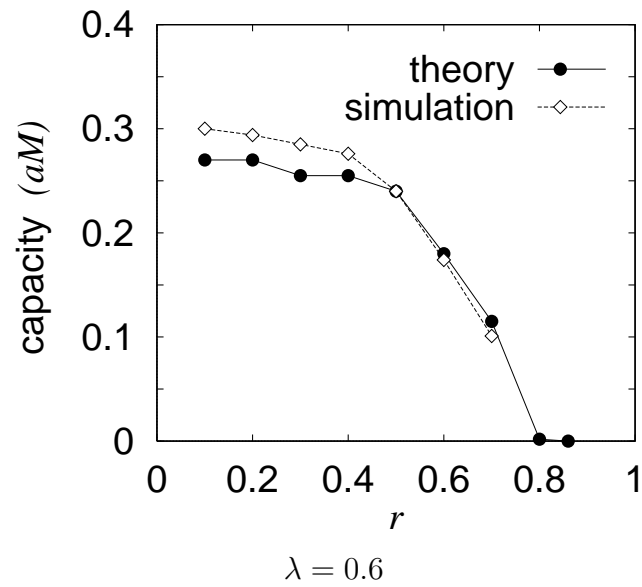


図 6.4: λ の違いによる記憶容量の変化 ($\lambda = 0.8, M = 3$)。パラメータ h は各相関での最適値を選んでいる。 λ が小さい方が概して記憶容量が高いが、相関 r が極めて大きいとは λ が大きい時にのみ選択的想起が実現できる。

6.3 部分反転法の統計神経力学

部分反転法の理論的解析には、既にいくつかの先行研究がある。一般の非単調ニューロンを用いたダイナミクス of 安定点解析としては Yoshizawa らの解析 [44] や Shiino と Fukai による SCSNA [41] などが行われている。部分反転法の安定点解析についてもその SCSNA が求められている [11]。一方、部分反転法のダイナミクスに関しては、Yanai と Amari がその第1 ステップの解析を行っている [43]。この Yanai と Amari が得た結果を漸化式の形にすると、統計神経力学の1 次近似モデルとなる。

この節ではこの統計神経力学の1 次近似モデルを用いて想起ダイナミクスを再現し、このモデルより得られる性質を実際のシミュレーションの結果と比較する。また、Okada により確立された統計神経力学の高次近似 [35] の考え方をうい、統計神経力学の2 次近似モデルを計算し、1 次近似、2 次近似、SCSNA、シミュレーションから得られる記憶容量を比較検討する。

6.3.1 部分反転法の統計神経力学

部分反転法の統計神経力学(1 次近似) は

$$m_{t+1} = \int Dz \langle \xi f_1(\xi m_t + \sigma_t z + \Gamma_t) \rangle_\xi \quad (6.31)$$

$$m_{t+2} = \int Dz \langle \xi f_0(\xi(m_t - m_{t+1}) + \rho_{t+1} z + \Gamma_{t+1}) \rangle_\xi \quad (6.32)$$

$$U_{t+1} = \frac{1}{\sigma_t} \int Dz z \langle \xi f_1(\xi m_t + \sigma_t z + \Gamma_t) \rangle_\xi \quad (6.33)$$

$$U_{t+2} = \frac{1}{\rho_{t+1}} \int Dz z z \langle \xi f_0(\xi(m_t - m_t) + \rho_{t+1} z + \Gamma_{t+1}) \rangle_\xi \quad (6.34)$$

$$S_{t,t+1} = a Q_{t,t+1} + U_{t+1} S_{t,t} \quad (6.35)$$

$$S_{t+1,t+1} = a Q_{t+1,t+1} + (U_{t+1})^2 S_{t,t} + 2a U_{t+1} Q_{t,t+1} \quad (6.36)$$

$$S_{t+2,t+2} = a Q_{t+2,t+2} + (U_{t+2})^2 (S_{t,t} + S_{t+1,t+1} - 2S_{t,t+1}) \\ + 2a U_{t+2} ((1 - U_{t+1}) Q_{t,t+2} - Q_{t+1,t+2}) \quad (6.37)$$

$$\rho_{t+1} = \sqrt{S_{t,t} + S_{t+1,t+1} - 2S_{t,t+1}} \quad (6.38)$$

$$\sigma_{t+2,t+2} = \sqrt{S_{t+2,t+2}} \quad (6.39)$$

$$Q_{t,t+1} = m_t m_{t+1} \quad (6.40)$$

$$Q_{t,t+2} = m_t m_{t+2} \quad (6.41)$$

$$Q_{t+1,t+1} = \int Dz \langle f_1(\xi m_t + \sigma_t z + \Gamma_t)^2 \rangle_\xi \quad (6.42)$$

$$Q_{t+2,t+2} = \int Dz \langle f_0(\xi(m_t - m_{t+1}) + \rho_{t+1} z + \Gamma_{t+1})^2 \rangle_\xi \quad (6.43)$$

$$Q_{t+1,t+2} = \int Dz_a \int Dz_b \int Dz_c \langle f_1(\xi m_t + \sigma_t (\sigma_t^{(\alpha)} z_a + \sigma_t^{(\beta)} z_b) + \Gamma_t) \\ \times f_0(\xi(m_t - m_{t+1}) + \text{sgn}(C_{t+1,t+2}) \rho_{t+1} (\sigma_t^{(\alpha)} z_a + \sigma_t^{(\beta)} z_c)) \rangle$$

$$+ \Gamma_{t+1}) > \xi \quad (6.44)$$

$$C_{t+1,t+2} = (1 - U_{t+1})S_{t,t} - aQ_{t,t+1} \quad (6.45)$$

$$\sigma_t^{(\alpha)} = \sqrt{\frac{|C_{t+1,t+2}|}{\sigma_t \rho_{t+1}}} \quad (6.46)$$

$$\sigma_t^{(\beta)} = \sqrt{1 - \frac{|C_{t+1,t+2}|}{\sigma_t \rho_{t+1}}} \quad (6.47)$$

$$Dz = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz \quad (6.48)$$

と求められる (付録 C)。ここで、関数 f_0, f_1 は、それぞれ $f_0(u) = \text{sgn}(u)$, $f_1(u) = \phi(u)$ に対応する。ノイズの項の平均値 Γ を正確に求めるのは困難であるが、 x_i に関して平均をとると

$$\Gamma_{t+1} = am_t U_{t+1} + U_{t+1} \Gamma_t \quad (6.49)$$

$$\Gamma_{t+2} = a(m_t - m_{t+1})U_{t+2} + U_{t+2}(\Gamma_t - \Gamma_{t+1}) \quad (6.50)$$

となる。このモデルは1次近似ではあるが、部分反転法の1単位ステップは2ステップの時間発展を含むので、近似の次数は通常ダイナミクスの2次近似に対応することになる。

従来の統計神経力学では、通常、ノイズの平均値 Γ は0であると仮定されている。SCSNAで得られる記憶パターン付近の安定点ではノイズの平均値が0になっていることから、想起に成功する場合はこの仮定は妥当であるといえる。 Γ を時間発展に組み込んだモデル [13] も検討されているが、このモデルでは時間が進むに連れて Γ が徐々に上昇し、実際のシミュレーションより系が過剰に安定化されてしまう。その結果、記憶容量を実際より高く算出してしまうことになる。

しかしながら、部分反転法の場合、 Γ の取り扱いについて安易に通常ダイナミクスの場合を踏襲するのは危険である。なぜなら、SCSNAの安定点では、 a が大きい場合、 Γ_{t+1} が無視できない大きい値をもつからである。一方、 Γ_{t+2} の方はあまり大きな値を持たない。そこで、以下の議論では、 $\Gamma_{t+2} = 0$, $\Gamma_{t+1} = am_t U_{t+1}$ を仮定したモデルを主に考えることにする。

6.3.2 統計神経力学の想起過程

まず、理論をシミュレーションの結果と比較参照するため、図 6.5 にシミュレーションで特徴的な振る舞いを示す3つの場合を示しておく。図 6.5(a) に示したように、 $a = 0.20$ の時は記憶パターン付近のパターンは完全に記憶パターンに引き込まれる。図 6.5(b) の $a = 0.28$ では、記憶パターン方向へ引き込みは起こるが、完全に収束せずに記憶パターン付近で不規則に振動する。図 6.5(c) の $a = 0.36$ の場合は、記憶パターンは完全に不安定化されている。

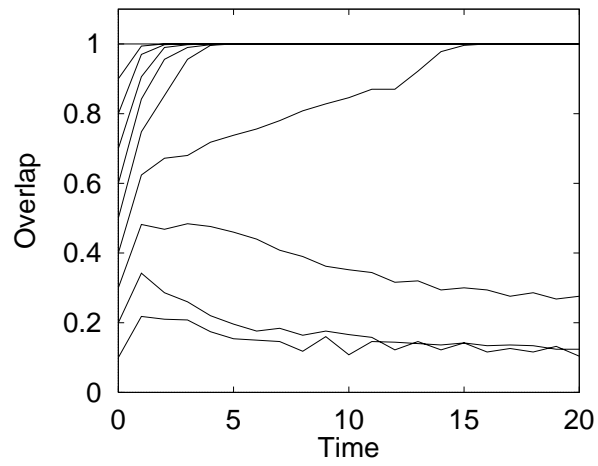
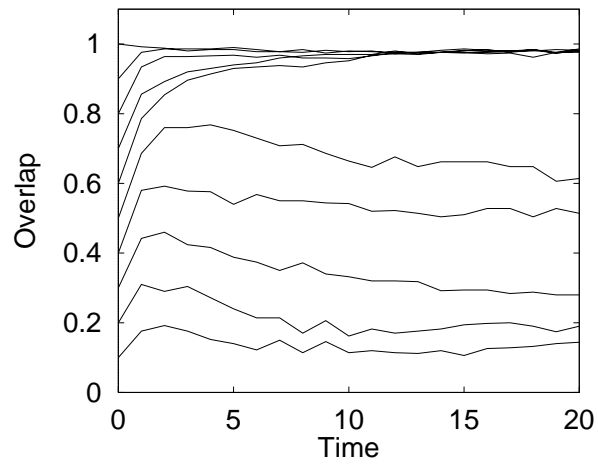
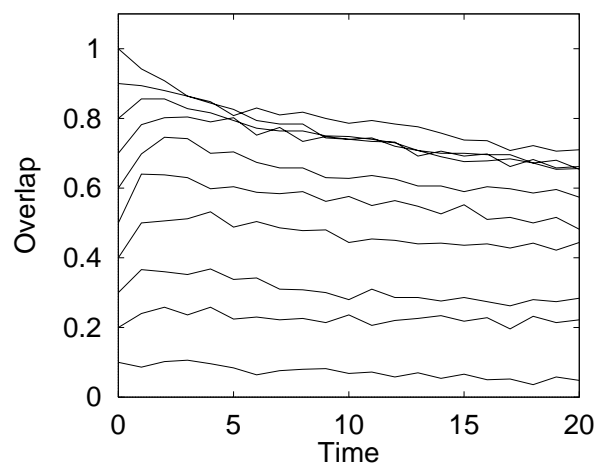
(a) $a = 0.20$ (b) $a = 0.28$ (c) $c = 0.36$

図 6.5: 部分反転法の想起ダイナミクスのシミュレーション ($N = 1000$, $h = 1.5$, $\lambda = 1.0$)。

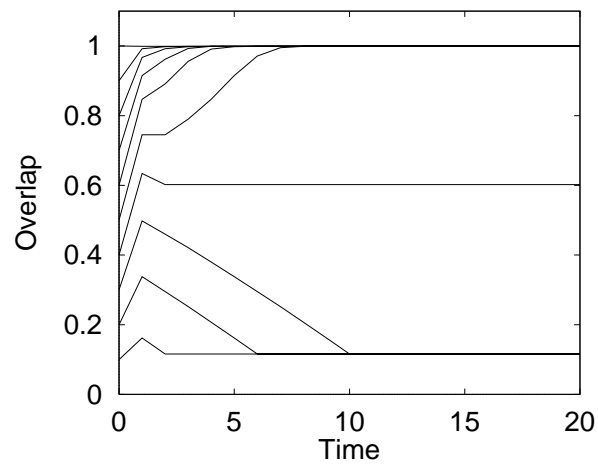
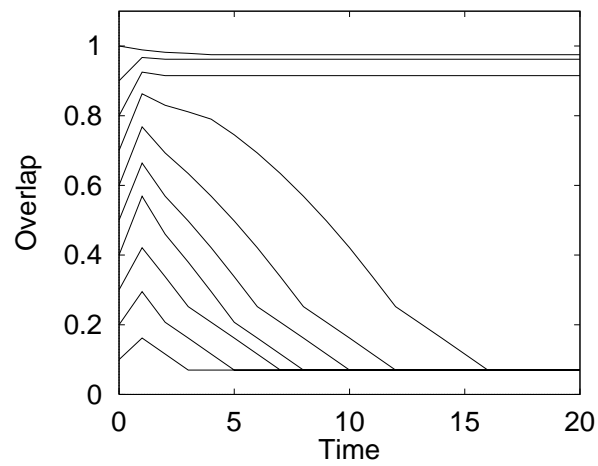
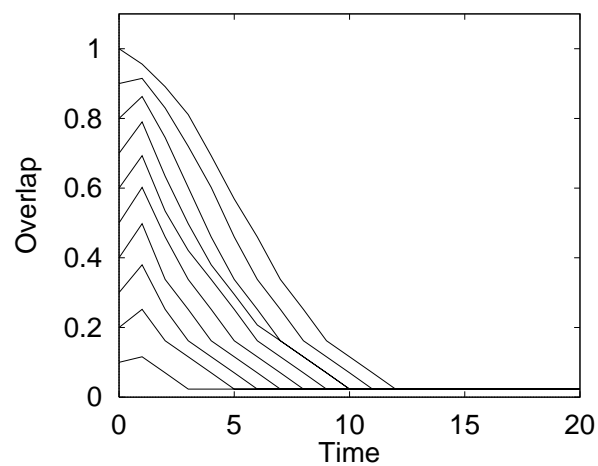
(a) $a = 0.20$ (b) $a = 0.28$ (c) $c = 0.36$

図 6.6: 部分反転法の想起ダイナミクスの統計神経力学1次近似モデル ($N = 1000$, $h = 1.5$, $\lambda = 1.0$)。

次に、前節で求めた統計神経力学において、 $\Gamma_{t+2} = 0, \Gamma_{t+1} = am_t U_{t+1}$ を仮定した1次近似モデルの想起ダイナミクスを図6.6に示す。想起に成功する場合は実際のシミュレーションに近い挙動が得られるが、想起に失敗する場合は収束先と記憶パターンとのオーバーラップが実際よりも小さくなる。これは統計神経力学一般に存在する問題である。シミュレーションで $a = 0.28$ 前後の時に現れる記憶パターン付近での不規則な振動は統計神経力学では現れない。統計神経力学のダイナミクスは、少数のマクロパラメータから構成されるため、このような振動状態の表現はできない。

上で述べたように、通常のダイナミクスに関する統計神経力学では、 Γ_{t+2} を考慮すると過剰な安定化が起り、記憶容量が高めに算出される。しかし、部分反転法の場合は、 Γ_{t+2} を時間発展に取り入れても、 Γ_{t+2} は小さい値に留まるので、ここで示した Γ_{t+2} を除去したモデルで得られる記憶容量に近い値が得られる。

6.3.3 各理論モデルの記憶容量の比較

前節では統計神経力学の1次近似を考えたが、このモデルで得られた記憶容量を、シミュレーション、SCSNA、そして統計神経力学の2次近似(計算結果略)と比較する(表6.1)。

シミュレーションにおいては、記憶パターンに収束する場合と記憶パターン付近で不規則に振動する場合があるので、明確な記憶容量は定義しにくい。が、 $a = 0.24$ 前後で不規則振動が増え始め、 $a = 0.29$ 前後でその記憶パターン付近の不規則振動も不安定になる。

SCSNAでは、記憶容量は $0.24N$ と計算され、シミュレーションの記憶容量の下限程度の値が得られる。SCSNAは安定点解析であることを考えると、記憶パターンが安定点となる限界の方が記憶容量として見積もられるのは妥当な結果である。

統計神経力学の場合、 Γ_{t+1} を考慮しないモデルでは記憶容量が実際のシミュレーションよりもかなり大きめに算出される。上で述べた通り、部分反転法のSCSNAでは安定点でも Γ_{t+1} は無視できない値を持つ。記憶容量の不一致はこれを切り捨てることにより生じていると予想される。実際、ノイズの平均値の項 Γ_{t+1} を考慮したモデルでは、シミュレーションに近い記憶容量が得られる。記憶パターンの想起に成功する時の安定点では、 $\Gamma_{t+1} = am_t U_{t+1}$ はSCSNAの安定点における Γ_{t+1} の良い近似になっている。また、この値はシミュレーションにおいて Γ_{t+1} に相当する量とも良く合致している。よって、少なくとも記憶パターン付近においては、 $\Gamma_{t+1} = am_t U_{t+1}$ を仮定したモデルの方が近似として適当であるといえる。

Γ_{t+1} を考慮した統計神経力学の場合、1次近似では記憶容量がSCSNAよりは大きめに計算され、付録に挙げた2次近似ではSCSNAに近い容量が得られる。これは、統計神経力学の近似次数が高くなるほどSCSNAに近づくという理論的事実[35]と合致している。

モデル	記憶容量
シミュレーション	0.24-0.29
SCSNA	0.24
統計神経力学1 次近似 ($\Gamma_{t+1} = 0$)	0.38
統計神経力学1 次近似 ($\Gamma_{t+1} \neq 0$)	0.28
統計神経力学2 次近似 ($\Gamma_{t+1} = 0$)	0.32
統計神経力学2 次近似 ($\Gamma_{t+1} \neq 0$)	0.25

表 6.1: 部分反転法のシミュレーションと各理論モデルの記憶容量の比較 ($h = 1.5, \lambda = 1.0$)。

6.3.4 引き込み領域

統計神経力学の時間発展方程式を得ることで、想起の引き込み領域を理論的に求めることが可能になる。上で導出した統計神経力学の1次近似モデルの引き込み領域をシミュレーションで得られた引き込み領域と比較したものを図6.7に示す。このように、記憶率 a が低い領域では統計神経力学の1次近似が精度の良い近似になっていることが分かる。しかし、 a が大きいときは Γ_{t+1} を考慮したモデルでは引き込み領域が小さく算出され、逆に Γ_{t+1} を考慮しないモデルでは引き込み領域が大きめに算出される。これは、ノイズの平均項を全てのニューロンについて平均化した $\Gamma_{t+1} = am_t U_{t+1}$ の評価が、記憶パターン付近では良い近似となっているものの、記憶パターンから離れたところでは、近似として不十分になっていることによると推測される。図6.7から、記憶パターンから遠い場所では、ノイズの平均項の効果が0と $am_t U_{t+1}$ の中間程度になっていると考えられる。

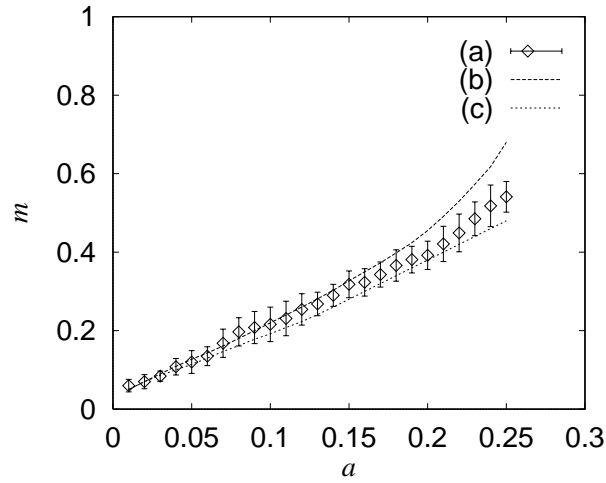


図 6.7: シミュレーションと統計神経力学から得られた引き込み領域の比較 ($h = 1.5, \lambda = 1.0$): (a) シミュレーション; (b) Γ_{t+1} を考慮した統計神経力学1次近似; (c) $\Gamma_{t+1} = 0$ と仮定した統計神経力学1次近似。

6.4 ニューロウインドウの統計神経力学

この節では、前節で議論した部分反転法の統計神経力学をもとに、ニューロウインドウ法の統計神経力学を求める。

ニューロウインドウ連想記憶は、相関のあるパターンを記憶した系に対し部分反転法を用いる。そこで、前節の部分反転法の統計神経力学と相関を持つパターンを記憶した場合の統計神経力学 [36] に基づき、ニューロウインドウ連想記憶の統計神経力学を計算すると、

$$m_{t+1}^{\mu} = \int Dz \langle \xi^{\mu} f_1 \left(\sum_{\mu \in A_1} \xi^{\mu} m_t^{\mu} + \sigma_t z + \Gamma_t \right) \rangle_{\xi} \quad (6.51)$$

$$m_{t+2}^{\mu} = \int Dz \langle \xi^{\mu} f_0 \left(\sum_{\mu \in A_1} \xi^{\mu} (m_t^{\mu} - m_{t+1}^{\mu}) + \rho_{t+1} z + \Gamma_{t+1} \right) \rangle_{\xi} \quad (6.52)$$

$$U_{t+1} = \frac{1}{\sigma_t} \int Dz z \langle f_1 \left(\sum_{\mu \in A_1} \xi^{\mu} m_t^{\mu} + \sigma_t z \right) \rangle_{\xi} \quad (6.53)$$

$$U_{t+2} = \frac{1}{\rho_{t+1}} \int Dz z \langle f_0 \left(\sum_{\mu \in A_1} \xi^{\mu} (m_t^{\mu} - m_{t+1}^{\mu}) + \rho_{t+1} z \right) \rangle_{\xi} \quad (6.54)$$

$$S_{t,t+1}^{(\kappa)} = a\lambda_\kappa^2 Q_{t,t+1} + \lambda_\kappa U_{t+1} S_{t,t}^{(\kappa)} \quad (6.55)$$

$$S_{t,t+1} = \sum_{\kappa \in A} S_{t,t+1}^{(\kappa)} \quad (6.56)$$

$$S_{t+1,t+1}^{(\kappa)} = a\lambda_\kappa^2 (Q_{t+1,t+1} + 2\lambda_\kappa U_{t+1} Q_{t,t+1}) + \lambda_\kappa^2 U_{t+1}^2 S_{t,t}^{(\kappa)} \quad (6.57)$$

$$S_{t+1,t+1} = \sum_{\kappa \in A} S_{t+1,t+1}^{(\kappa)} \quad (6.58)$$

$$\begin{aligned} S_{t+2,t+2}^{(\kappa)} = & a\lambda_\kappa^2 (Q_{t+2,t+2} + \lambda_\kappa^2 U_{t+2}^2 Q_{t+1,t+1} - 2\lambda_\kappa U_{t+2} Q_{t+1,t+2} \\ & + 2\lambda_\kappa U_{t+2} (1 - \lambda_\kappa U_{t+1}) Q_{t,t+2} - 2\lambda_\kappa^2 U_{t+2}^2 (1 - \lambda_\kappa U_{t+1}) Q_{t,t+1}) \\ & + \lambda_\kappa^2 U_{t+2}^2 (1 - \lambda_\kappa U_{t+1})^2 S_{t,t}^{(\kappa)} \end{aligned} \quad (6.59)$$

$$S_{t+2,t+2} = \sum_{\kappa \in A} S_{t+2,t+2}^{(\kappa)} \quad (6.60)$$

$$\rho_{t+1} = \sqrt{S_{t,t} + S_{t+1,t+1} - 2S_{t,t+1}} \quad (6.61)$$

$$Q_{t,t+1} = m_t^1 m_{t+1}^1 \quad (6.62)$$

$$\sigma_{t+2} = \sqrt{S_{t+2,t+2}} \quad (6.63)$$

$$Q_{t,t+2} = m_t^1 m_{t+2}^1 \quad (6.64)$$

$$Q_{t+1,t+1} = \int dz \langle f_1(\sum_{\mu \in A_1} \xi^\mu m_t^\mu + \sigma_t z)^2 \rangle_\xi \quad (6.65)$$

$$Q_{t+2,t+2} = \int dz \langle f_0(\sum_{\mu \in A_1} \xi^\mu (m_t^\mu - m_{t+1}^\mu) + \rho_{t+1} z)^2 \rangle_\xi \quad (6.66)$$

$$\begin{aligned} Q_{t+1,t+2} = & \int dz_a \int dz_b \int dz_c \langle f_1(\sum_{\mu \in A_1} \xi^\mu m_t^\mu + \sigma_t (\sigma_t^{(\alpha)} z_a + \sigma_t^{(\beta)} z_b) \\ & + \Gamma_t) \times f_0(\sum_{\mu \in A_1} \xi^\mu (m_t^\mu - m_{t+1}^\mu) \\ & + \text{sgn}(C_{t,t+1}) \rho_{t+1} (\sigma_t^{(\alpha)} z_a + \sigma_t^{(\beta)} z_c) + \Gamma_{t+1}) \rangle_\xi \end{aligned} \quad (6.67)$$

$$C_{t,t+1}^{(\kappa)} = a\lambda_\kappa^2 (1 - \lambda_\kappa U_{t+1}) Q_{t,t+1}^{(\kappa)} + \lambda_\kappa U_{t+1} S_{t,t} \quad (6.68)$$

$$C_{t,t+1} = \sum_{\kappa \in A} C_{t,t+1}^{(\kappa)} \quad (6.69)$$

$$\sigma_t^{(\alpha)} = \sqrt{\frac{|C_{t,t+1}|}{\sigma_t \rho_{t+1}}} \quad (6.70)$$

$$\sigma_t^{(\beta)} = \sqrt{1 - \frac{|C_{t,t+1}|}{\sigma_t \rho_{t+1}}} \quad (6.71)$$

$$Dz = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz \quad (6.72)$$

という漸化式が得られる (付録 D)。ここで、今着目するパターンを ξ^1 、それと相関のあるパターン群 (ξ^1 も含む) を $\xi^\mu (\mu \in A_1)$ とし、それらと \boldsymbol{x} の相関を m^μ としている。更

に、 $\lambda_\kappa (\kappa \in A)$ は、

$$k_{ij} = \begin{cases} 1, & \text{if } i = j \\ r^2, & \text{if } i \neq j \end{cases} \quad (6.73)$$

を満たす $M \times M$ 行列 $K = [k_{ij}]$ の固有値である。

以上で得られたニューロウインドウ連想記憶の統計神経力学を使って、数値実験の結果が理論的に再現されるかどうかを以下の議論で検証する。

まず、ニューロウインドウ連想記憶の特徴は、パラメータ h を変化させて記憶パターンと概念パターンを選択的想起する点にあるので、アトラクタが移動する様子を捉えることにする。図 6.8 に統計神経力学で異なる h を使った想起過程のオーバーラップの変化を求めた結果を示す ($a = 0.03, M = 3, r = 0.7, \lambda = 0.8$)。このように、得られた理論によって、 $h = 1.6$ で記憶パターンに引き込まれ、 $h = 2.4$ で概念パターンに引き込まれる様子が再現されており、想起までに必要な時間等についてもシミュレーションと一致した特徴が得られている。

次に、ニューロウインドウ連想記憶の統計神経力学で新たに得られた引き込み領域の理論値を示す。図 6.9 と図 6.10 はそれぞれ、記憶パターンを安定化するパラメータ ($h = 1.6, \lambda = 0.8$) を使った時と概念パターンを安定化するパラメータ ($h = 2.4, \lambda = 0.8$) を使った時の記憶／概念パターンの安定性と引き込み領域の実験値と理論値を比較した結果である ($M = 3, r = 0.6$ と $M = 3, r = 0.7$)。ここで、前節で述べた通り、部分反転法の統計神経力学では、記憶の安定性を調べる場合はノイズの平均項 Γ を一部考慮すると良い近似になるので、ここでも安定性の議論では前節で使った近似 ($\Gamma_t = 0, \Gamma_{t+1} = am_t U_{t+1}$) を用いている。引き込み領域については、実験値は Γ_{t+1} を考慮する理論値と $\Gamma_{t+1} = 0$ とする理論値の中間に位置し、特に記憶率 a が大きい場合は Γ_{t+1} を考慮するモデルは引き込み領域を狭く見積もるという結果が部分反転法の統計神経力学で得られている。そこで、引き込み領域の理論値については、 Γ を考慮するモデルと無視するモデルの両モデルの理論値を示している。図 6.9、図 6.10 から分かるように、安定性の理論値は実験結果を良く説明している。また、引き込み領域についても、部分反転法の統計神経力学で得られた結果と同様の結果が得られている。

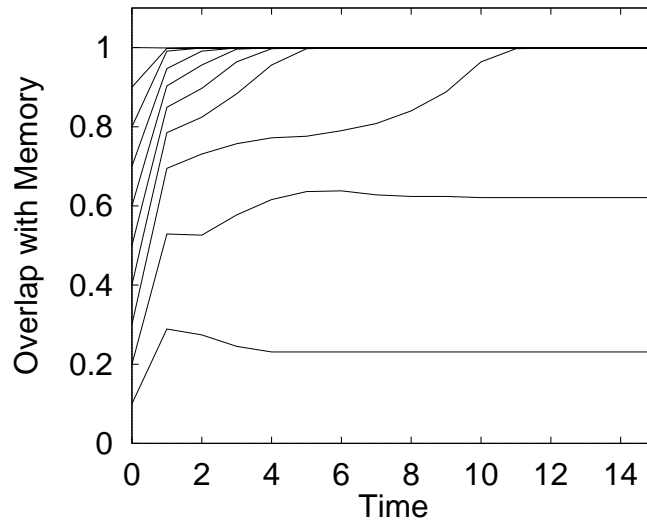
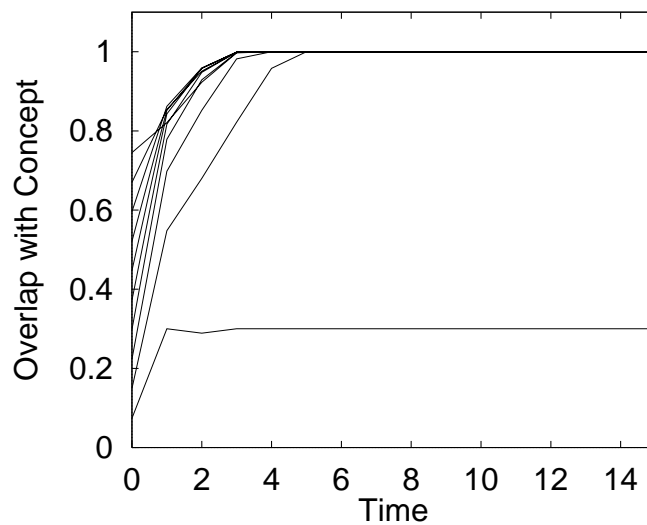
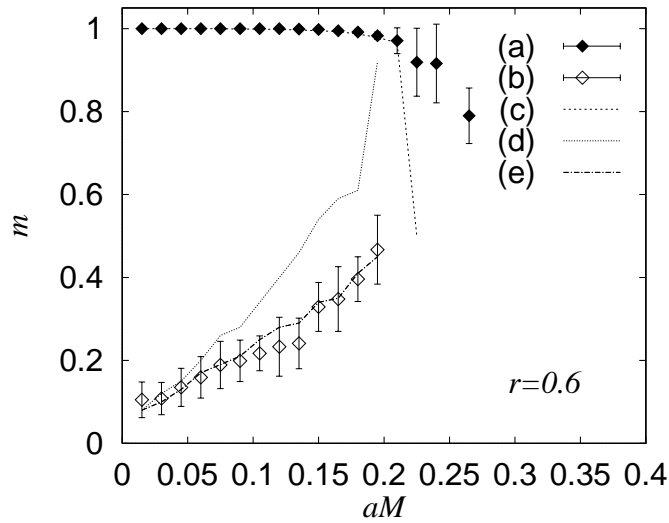
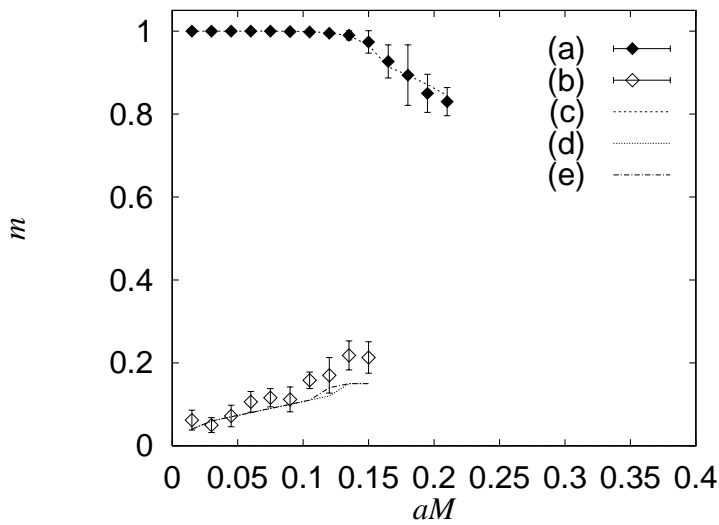
(A) 記憶パターンの想起過程 ($h = 1.6, \lambda = 0.8$)(B) 概念パターンの想起過程 ($h = 2.4, \lambda = 0.8$)

図 6.8: 記憶/概念パターンの想起過程 ($a = 0.03, M = 3, r = 0.7$)。小さい h では記憶パターンへの引き込みが起こり、大きい h では概念パターンへの引き込みが起こる。

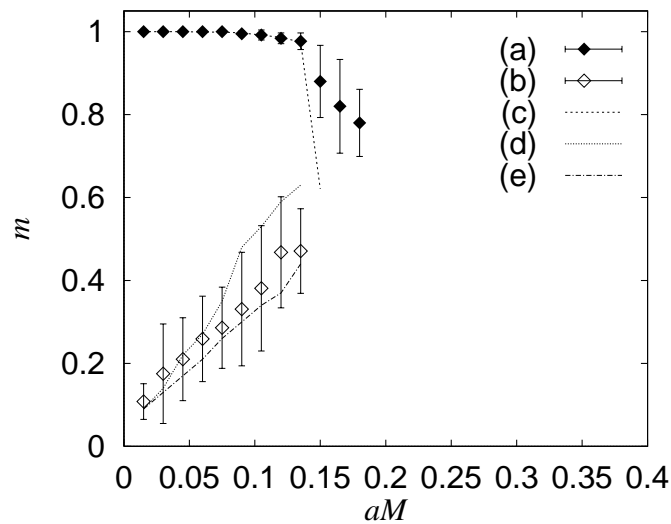


(A) 記憶パターン ($h = 1.6, \lambda = 0.8$)

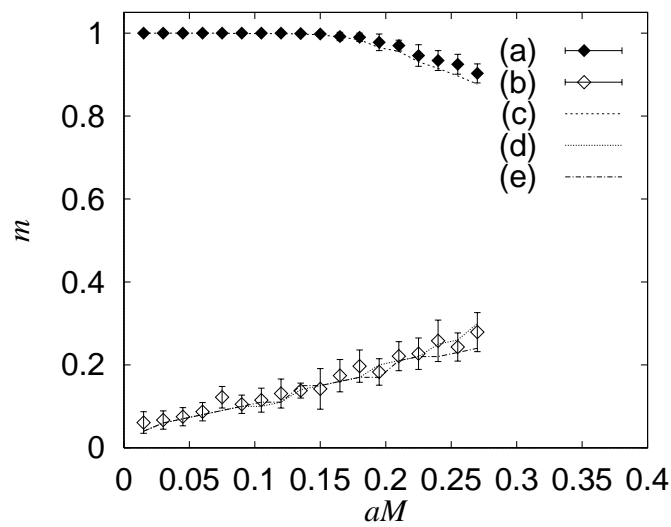


(B) 概念パターン ($h = 2.2, \lambda = 0.8$)

図 6.9: $M = 3, r = 0.6$ における記憶パターンと概念パターンの収束先(安定性)と引き込み領域の理論値と実験結果: (a) 安定性の実験値; (b) 引き込み領域の実験値; (c) 安定性の理論値 (Γ 考慮); (d) 引き込み領域の理論値 (Γ 考慮); (e) 引き込み領域の理論値 (Γ 無視)。



(A) 記憶パターン ($h = 1.6, \lambda = 0.8$)



(B) 概念パターン ($h = 2.4, \lambda = 0.8$)

図 6.10: $M = 3, r = 0.7$ における記憶パターンと概念パターンの収束先(安定性)と引き込み領域の理論値と実験結果: (a) 安定性の実験値; (b) 引き込み領域の実験値; (c) 安定性の理論値 (Γ 考慮); (d) 引き込み領域の理論値 (Γ 考慮); (e) 引き込み領域の理論値 (Γ 無視)。

6.5 まとめ

本章では、ニューロウインドウ法について統計手法を用いた解析を行った。そこで得られた理論値と実験値の比較を行った。得られた理論はニューロウインドウ法の選択的想起におけるパラメータの境界値や記憶容量等の定量値を与えることに成功した。

本章では、最も簡単な二段の相関を持つパターン群を記憶した系でニューロウインドウ法による選択的想起を行う場合について議論を行った。しかし、前章で述べた通り、ニューロウインドウ法には他にも様々な応用がある。今後はそれらの応用モデルについての定量的な解析を進める必要がある。また、ニューロウインドウ連想記憶に限らず、ランダムパターンにおける相関が産み出す固有値の分布を考慮した、記憶パターンと偽記憶パターンの想起についての考察なども今後の課題として挙げられる。

第 7 章： 連続時間ダイナミクス

前章までは、離散時間ダイナミクスを持つ系を対象を限って議論を展開してきた。本章は、連続時間ダイナミクスを持つ系を考え、その性質について議論を行う。まず、第1節で連続時間ニューロダイナミクスを定式化する。次に、第2節において、対称な結合荷重を持つ連続時間系自己連想記憶を取り上げ、連続時間ニューロダイナミクスの性質を調べる。更に、第5章で導入したニューロウィンドウ法を連続時間系に拡張する。第3節では、連続時間系相互相関連想記憶を取り上げ、系の解析とニューロウィンドウ法の応用を行う。第4節では、荷重行列の非対称性が強いランダムネットワークを取り上げ、連続時間系において非対称荷重が作り出す力学系を解析する。

7.1 連続時間ニューロダイナミクス

第2章で述べた通り、連続時間型ニューラルネットワークにおけるダイナミクスは

$$\tau \frac{du_i}{dt} = -u_i + \sum_{j=1}^N w_{ij} x_j \quad (7.1)$$

$$x_j = f(u_j) \quad (7.2)$$

で与えられる。

離散時間ダイナミクスにおいては、荷重行列 $W = [w_{ij}]$ が対称行列の場合、エネルギー関数は

$$E(\mathbf{x}(t)) = -\frac{1}{2} \mathbf{x}^T(t+1) W \mathbf{x}(t) \quad (7.3)$$

で与えられる(第3章参照)。また、D関数

$$D(\mathbf{x}(t)) = \mathbf{x}^T(t) W^T W \mathbf{x}(t) \quad (7.4)$$

が想起過程においてほぼ単調に増加する(但し、非線形の効果で減少することもある)ことから、 $-D(\mathbf{x}(t))$ がエネルギー関数に近い性質を持つ。 $W^T W$ の二次形式が増加することは、状態ベクトル \mathbf{x} を構成するベクトル成分のうち、行列 $W^T W$ の大きい固有値をもつ固有ベクトル成分が増えていくことを意味し、それは W の 絶対値の大きい固有値 をもつ固有ベクトル成分が増加することに対応している。

一方、連続時間型ダイナミクスについては既に多くの研究がされており [18]、荷重行列が対称の場合、エネルギー関数は

$$E(\mathbf{x}(t)) = -\frac{1}{2} \mathbf{x}^T(t) W \mathbf{x}(t) \quad (7.5)$$

で与えられることが知られている。このエネルギーが単調減少であるので、 W の二次形式 $\mathbf{x}^T W \mathbf{x}$ は単調増加となる。 W の二次形式が増加することは、状態ベクトル \mathbf{x} を構成するベクトル成分のうち、 W の 正の大きい固有値 をもつ固有ベクトル成分が増えていくことを意味している。これが離散時間ダイナミクスとの特徴的違いの一つである。

上の連続時間ダイナミクスは微分方程式で与えられるが、実際にデジタル計算機でシミュレーションをする場合は差分方程式

$$u_i(t + \tau\Delta) = (1 - \Delta)u_i(t) + \Delta \sum_{j=1}^N w_{ij}x_j(t) \quad (7.6)$$

に変換して計算を行うことになる。ここで、 Δ は $\frac{1}{\tau}dt$ に相当する。この離散化法はオイラー差分を使った最も単純な方法で、連続時間ダイナミクスをより正確に近似するにはルンゲ-クッタ法等を適用すべきであるが、ここでは、離散時間モデルとの理論的対比を容易にするため、オイラー法を前提に議論を進める。

7.2 連続時間自己連想記憶

この節においては、前節で述べた連続時間ニューロダイナミクスを用いた自己相関連想記憶について議論を行う。まず、7.2.1 で連続時間自己相関連想記憶の一般的性質について述べる。次に、7.2.2 で連続時間自己相関連想記憶にニューロウインドウ法を導入する。最後に、7.2.3 で連続時間系に対して有限系解析を適用し、離散時間系との比較を行う。

7.2.1 連続時間自己相関連想記憶

連続時間自己相関連想記憶においては、荷重行列は離散時間の場合と同じものを用いる。そして、想起ダイナミクスとして前節で定義した連続時間ダイナミクスを考える。連続時間ダイナミクスを用いた場合の想起の様子 (記憶率 $a = 0.1, a = 0.2$) を図 7.1 に示す。但し、ここで、各ニューロンの内部状態の初期値 $u_i(0)$ は

$$u_i(0) = x_i(0) \quad (7.7)$$

で与えている。この図から、連続時間モデルにおいても、記憶パターンの想起 (図 7.1 (a))、記憶容量超過時の偽記憶 (図 7.1 (b)) といった離散時間モデルと同様の挙動が現れることが分かる。また、状態遷移の様子を幾何的に見るため、固有ベクトル成分の変化を計算した結果を図 7.2 に示す。このように、記憶容量を超えてパターンを記憶した時に生じる偽記憶は、ここでも大きな固有値を持つ固有ベクトル方向に存在していることが分かる。

連続時間ニューロダイナミクスにおいて、非単調ニューロン (図 7.3) を用いたダイナミクス

$$\tau \frac{du_i}{dt} = -u_i + \sum_{j=1}^N w_{ij}x_j \quad (7.8)$$

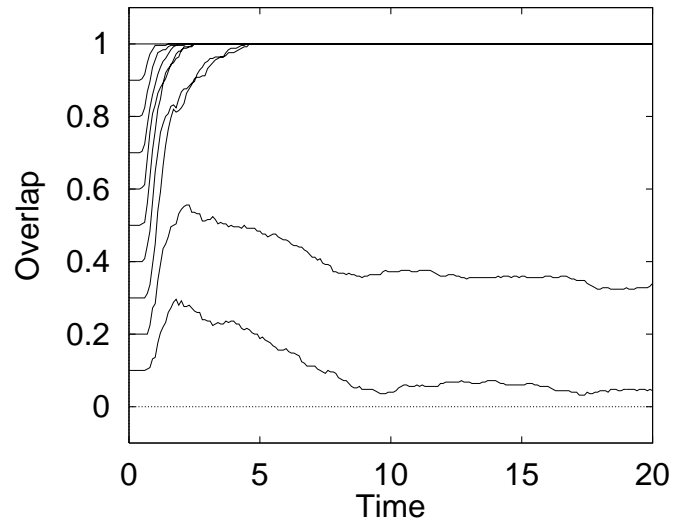
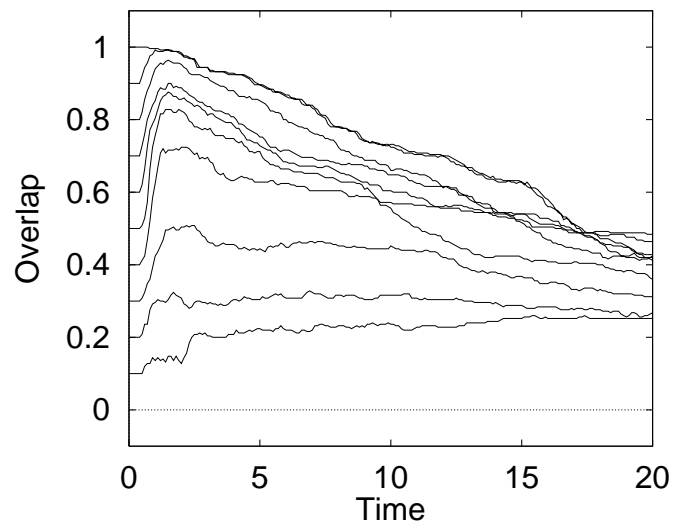
(a) $a = 0.1$ (b) $a = 0.2$

図 7.1: 連続時間自己相関連想記憶の想起過程におけるオーバーラップの変化 ($N = 500, \tau = 1, \Delta = 0.1$)。

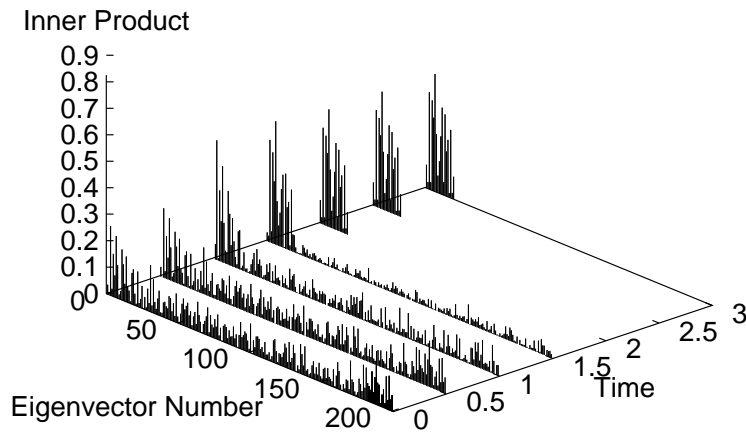
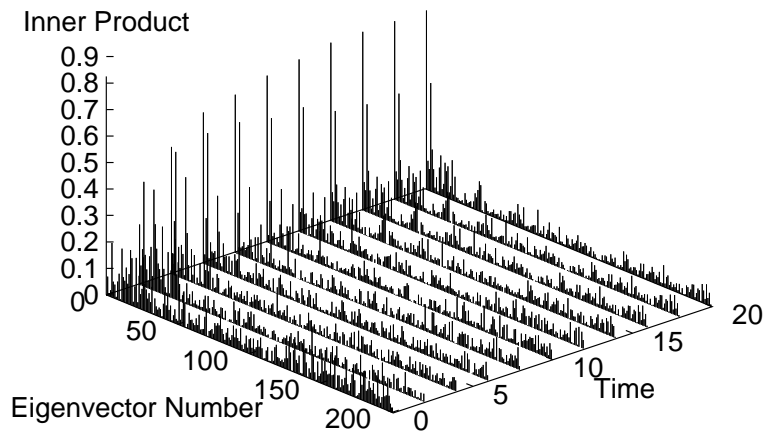
(a) $a = 0.1$ (b) $a = 0.2$

図 7.2: 連続時間自己相関連想記憶の想起過程における固有ベクトル成分の変化 ($N = 200, \tau = 1, \Delta = 0.1, m(0) = 0.5$)。

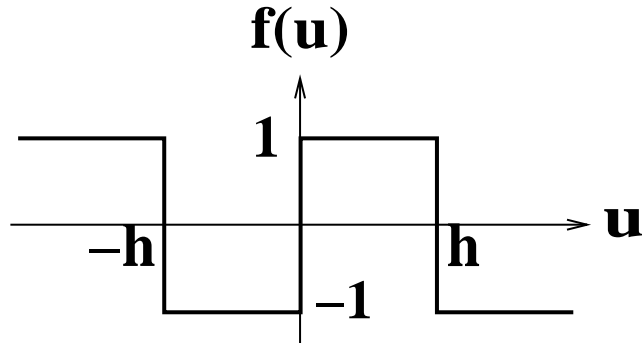


図 7.3: 連続時間モデルにおける非単調ニューロンの活性化関数 $f(u)$ 。

$$x_j = f(u_j) \quad (7.9)$$

$$f(u) = \begin{cases} 1, & \text{if } u < -h \\ -1, & \text{if } -h \leq u < 0 \\ 1, & \text{if } 0 \leq u < h \\ -1, & \text{if } h \leq u \end{cases} \quad (7.10)$$

が自己相関連想記憶の記憶容量を増大させることが知られている [29]。但し、ここで系の観測に関しては通常使われる符号関数のニューロンを用い、 $y_j = \text{sgn}(u_j)$ で得られる状態ベクトル \mathbf{y} を想起パターンとする。つまり、オーバーラップは

$$m^p = \frac{1}{N} \sum_{i=1}^N \text{sgn}(u_i) s_i^p \quad (7.11)$$

で定義している。このダイナミクスでは、内部状態が大きいニューロンの出力が反転されるので、部分反転法と同様に絶対値の大きい固有値を持つ固有ベクトル方向への流れを抑える効果が現れる。その結果、偽記憶への状態遷移は抑えられ、記憶パターンが安定化される。よって、非単調ニューロンとしては、内部状態が大きいニューロンの出力が弱められるような関数であれば、上記に挙げた関数に限らず記憶容量増大に有効に働く。

実際、非単調ニューロンを用いて想起を行った場合のオーバーラップと固有ベクトル成分の変化を図 7.4 と図 7.5 に示す。この様に、非単調ニューロンによって、単調ニューロンでは想起が不可能な記憶率でも、記憶パターンが想起されていることが分かる。非単調ニューロンを使った連続時間モデルでは、離散時間モデルの部分反転法に比べ記憶容量の増大幅が大きく、パラメータを適切に選ぶと $0.4N$ 程度の記憶容量が達成できることが知られている [29, 44]。

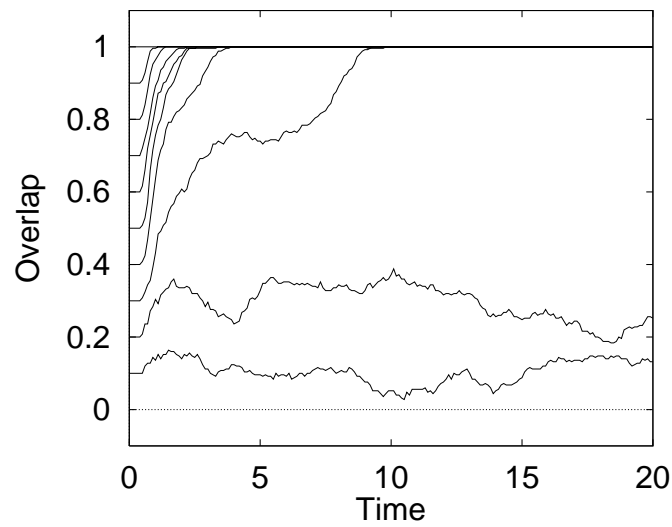


図 7.4: 非単調ニューロンを使った連続時間自己相関連想記憶の想起過程におけるオーバーラップの変化 ($a = 0.2, N = 500, h = 1.2, \tau = 1, \Delta = 0.1$)。

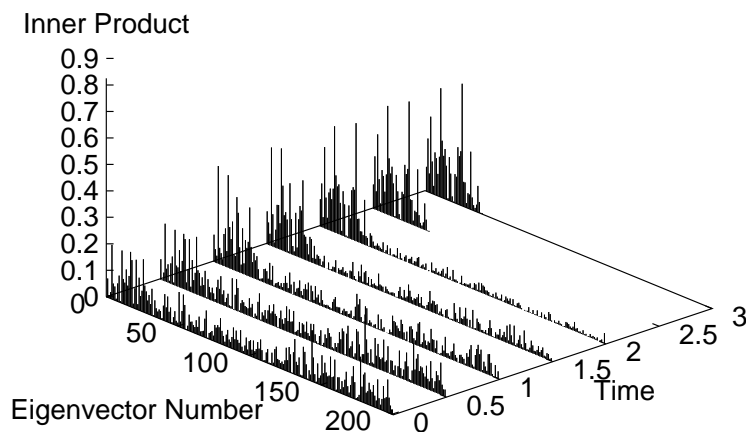


図 7.5: 非単調ニューロンを使った連続時間自己相関連想記憶の想起過程における固有ベクトル成分の変化 ($a = 0.2, N = 500, h = 1.2, \tau = 1, \Delta = 0.1, m(0) = 0.5$)。

7.2.2 連続時間ニューロウインドウ法

前小節で述べた通り、連続時間ダイナミクスにおいては、非単調ニューロンによって大きい固有値を持つ固有ベクトル方向への状態遷移が抑えられる。このことから、この非単調ニューロンのパラメータ h を調節し、窓の大きさを制御することで、連続時間ダイナミクスにおけるニューロウインドウ法が実現できると期待される。

そこで、第5章で論じた階層的相関構造

$$\frac{1}{N} E[\mathbf{s}^{(p_1, p_2, \dots, p_L)} \cdot \mathbf{s}^{(q_1, q_2, \dots, q_L)}] = \begin{cases} 1, & \text{if } p_1 = q_1, \dots, p_L = q_L \\ R^{L-l}, & \text{if } p_1 = q_1, \dots, p_l = q_l, p_{l+1} \neq q_{l+1} \quad (1 \leq l \leq L-1) \\ 0, & \text{if } p_1 \neq q_1 \end{cases} \quad (7.12)$$

を持つ記憶パターン $\mathbf{s}^{(p_1, p_2, \dots, p_L)}$ を考え、それらの自己相関を記憶した荷重行列を持つネットワークに対し、連続時間ニューロウインドウ法の適用を試みた。図7.6に、第5章でも例題として扱った $L=3, P_2=P_3=3, R=0.49$ のケースについて、非単調ニューロンのパラメータ h を変化させて想起した時の結果を示す。この図に示されている通り、連続時間モデルにおいても、非単調ニューロンの窓の大きさを調節することで、記憶パターンや各階層の概念パターンが選択的に安定化できることが分かる。

更に、第二概念パターンが安定化されるようにパラメータを設定した時に、記憶パターンに近いパターンから想起を始めた場合の想起過程を図7.7に示す。図のように、連続時間モデルでは、上位の概念パターンを想起する途中で記憶パターン付近に一時停留し、その後概念パターンへと引き込まれる現象が見られる。これは、離散時間モデルにはない特徴的な現象である。

以上の数値実験から、連続時間ニューロウインドウ法と離散時間ニューロウインドウ法の相違点として、連続時間モデルでは記憶パターンの引き込み領域が大きいこと、及び上位の概念パターンの想起において記憶や下位の概念の一時的想起状態を経由することが分かった。この相違が発生するメカニズムを調べるため、連続時間モデルと離散時間モデルの想起過程をD関数を用いて比較検討することにする。

まず、記憶パターンへの引き込みの過程について調べる(図7.8)。離散時間モデルでは、初期値が記憶パターンから離れている場合は、上位の概念パターンへと引き込まれていく。この過程におけるD関数の変化を見ると、時刻 $t=1$ で $D(\mathbf{x})$ の値が大きく跳ね上がっている。このことは、想起の第一ステップで、状態が既に上位の概念パターンが分布する下流まで流されてしまっていることを意味する。連続時間モデルにおいては、状態遷移がゆっくりと進むため、初期値が記憶パターンから離れていても、途中で記憶パターンの分布するバンドを経由する。そのため、記憶パターンのバンドを横切る時に、記憶パターンへの引き込みが可能となり、それが記憶パターンの引き込み領域が拡大につながっている。

次に、第二概念パターンを安定化した時の、記憶パターン付近からの想起について、D関数の変化をモニターする(図7.9)。ここでも、離散時間モデルにおいては、想起の第一ステップで上位の概念パターンの分布するバンドまで一気に流される。ところが、連

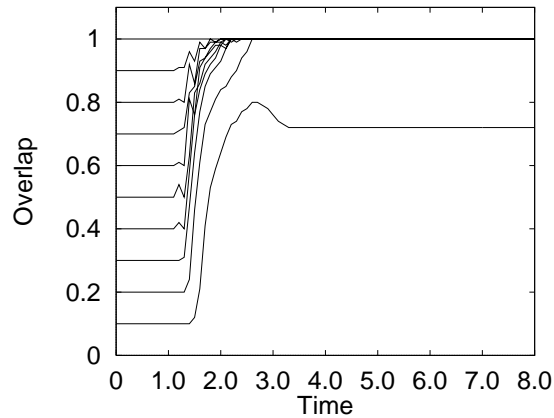
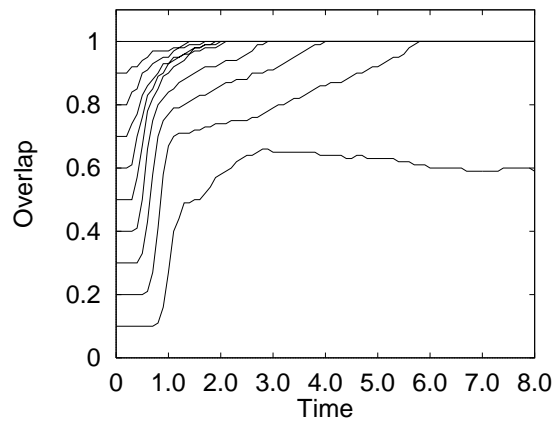
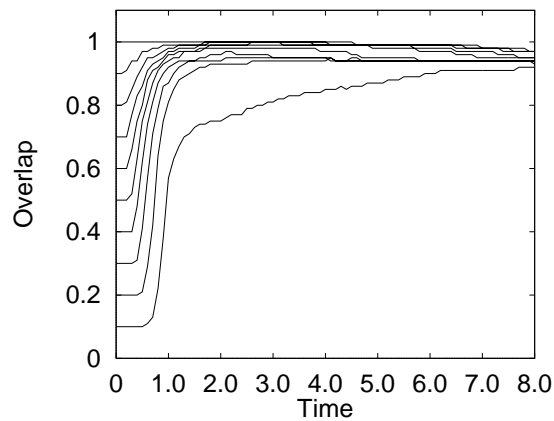
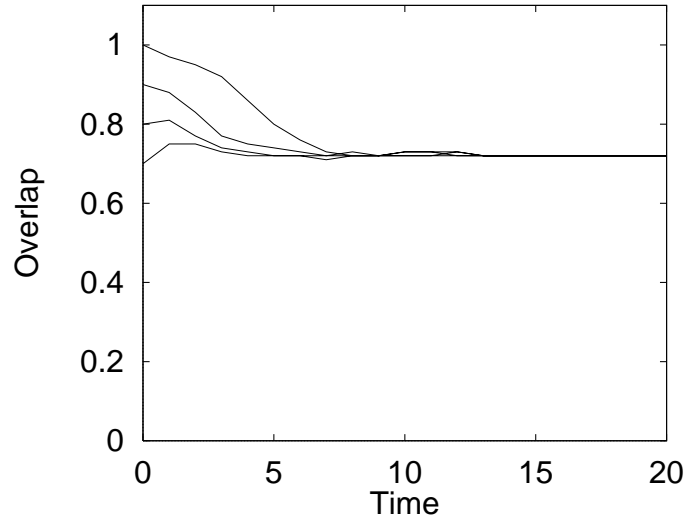
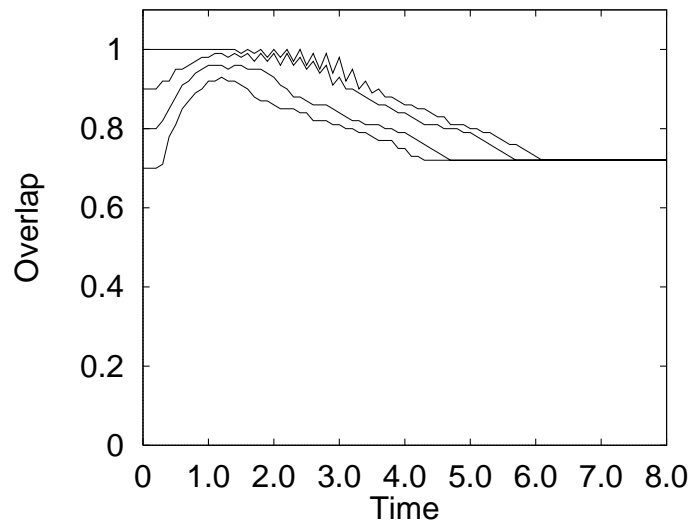
(a) 記憶パターンとのオーバーラップの変化 ($h = 0.5$)(b) 第二概念パターンとのオーバーラップの変化 ($h = 1.5$)(c) 第一概念パターンとのオーバーラップの変化 (単調ニューロン: $h = \infty$)

図 7.6: 連続時間モデルにおける階層概念の選択的想起 ($N = 1000, L = 3, Q_1 = 5, Q_2 = Q_3 = 3, R = 0.49$)。



(a) 離散時間モデル ($h = 2.9, \lambda = 1.0$)



(b) 連続時間モデル ($h = 1.5$)

図 7.7: 第二概念パターンを安定化させた時の記憶パターンに近いパターンからの想起過程における記憶パターンとのオーバーラップの変化 ($N = 1000, L = 3, Q_1 = 5, Q_2 = Q_3 = 3, R = 0.49$)。

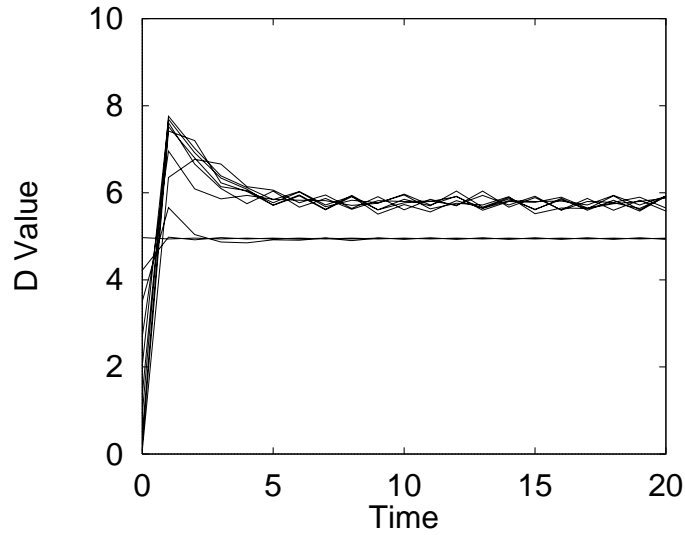
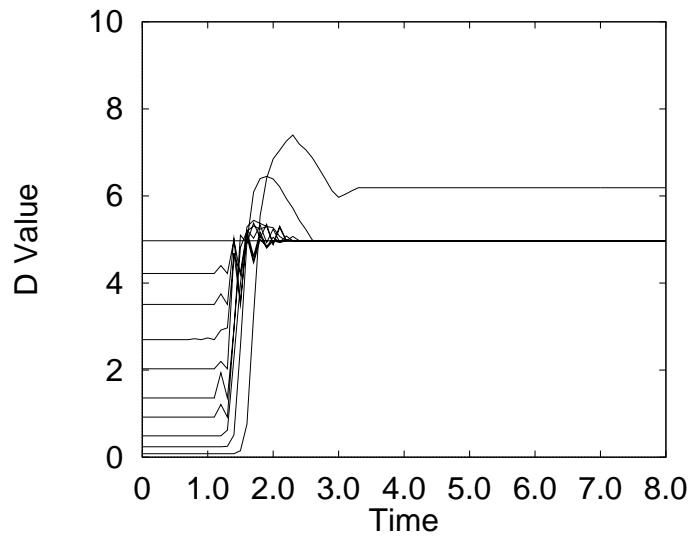
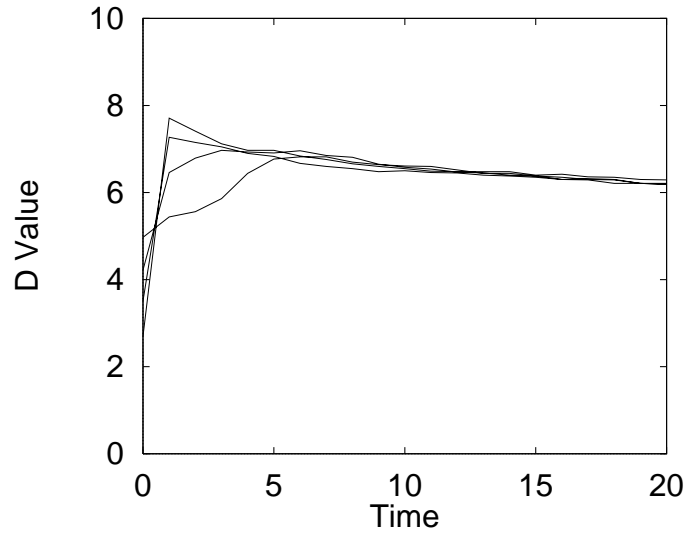
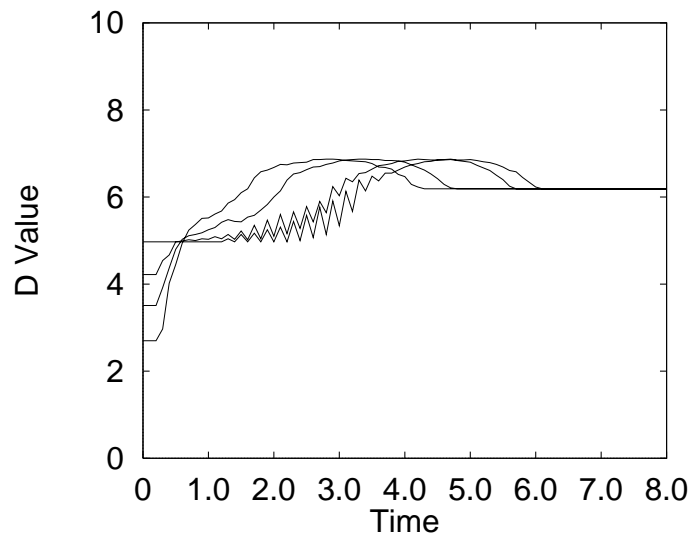
(a) 離散時間モデル ($h = 2.5, \lambda = 1.0$)(b) 連続時間モデル ($h = 0.5$)

図 7.8: 記憶パターンを安定化させた時の想起過程における D 関数の値の変化 ($N = 1000, L = 3, Q_1 = 5, Q_2 = Q_3 = 3, R = 0.49$)。



(a) 離散時間モデル ($h = 2.9, \lambda = 1.0$)



(b) 連続時間モデル ($h = 1.5$)

図 7.9: 第二概念パターンを安定化させた時の記憶パターンに近いパターンからの想起過程における D 関数の値の変化 ($N = 1000, L = 3, Q_1 = 5, Q_2 = Q_3 = 3, R = 0.49$)。

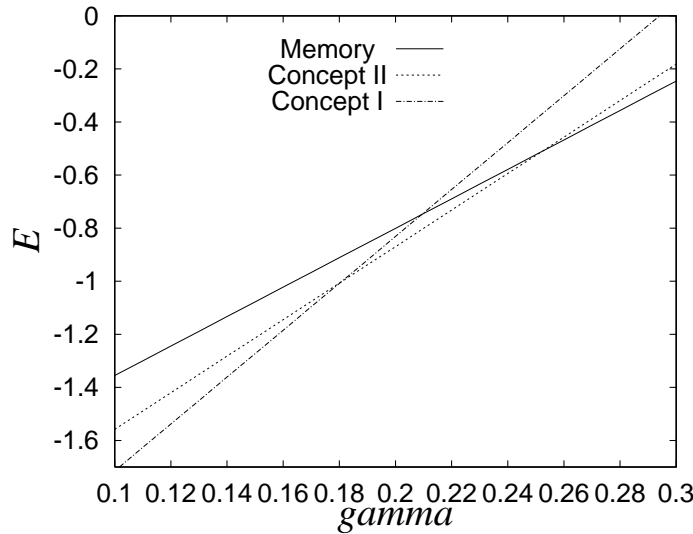


図 7.10: γ を変えたときの記憶／概念パターンのエネルギー。 γ を増やすと、最小エネルギー状態が、第一概念パターンから、第二概念パターン、記憶パターンへと変化していくことが分かる。

連続時間モデルにおいては、状態遷移が緩やかに進むため、記憶パターンの分布するバンドにたどり着いた時、一度記憶パターンの近くで一度停留し、その後、状態はより下流へと流れていくことになる。これが、記憶パターンに一旦想起した後、概念パターンを想起するように見える原因である。

尚、非単調ニューロンを用いた連続時間モデルが部分反転法よりも大きい記憶容量を持つことから類推される通り、ニューロウインドウ法においても、記憶パターン、概念パターンとも連続時間モデルは離散時間モデルよりも記憶容量が大きく、選択的想起が可能な条件の範囲が広い。

また、離散時間モデルの部分反転法の時と同様に連続時間モデルにおいても、非単調ニューロンを用いる代わりに荷重行列を

$$W = W - \gamma W^2 \quad (7.13)$$

と変更することでニューロウインドウ連想記憶を実現することができる。このことは、エネルギーの概念を用いて以下のように説明することが可能である。

本小節の議論で例題として用いた $L = 3, P_2 = P_3 = 3, R = 0.49$ の場合について、 γ を変化させたときに得られる記憶パターンと概念パターンのエネルギー

$$E(\mathbf{x}) = \mathbf{x}^T (W - \gamma W^2) \mathbf{x} \quad (7.14)$$

を比較すると図 7.10 のようにある。このように、 γ を大きくすると、より下位の概念が低エネルギー状態になるので、 γ を調節することで系の安定点を選択できることが分かる。その結果、各階層の記憶／概念パターンの選択的想起が実現されるのである。

7.2.3 有限系の解析

この小節では、連続時間モデルに対して、第6章で述べた有限系解析を行い、非単調ニューロンが選択的想起を可能にするメカニズムを追う。

ここでは、議論を簡単にするため、相関構造が最も単純な例として、親パターン s^0 と相関 r を持つ子パターン s^1, s^2, s^3 の3パターンだけを記憶した場合について解析する。以下、親パターンと同じ発火状態のニューロンを (+) 状態、反対の発火状態のニューロンを (-) 状態と記す。

第6章で述べた通り、全ニューロンは各記憶パターンでとる値によって分類することができる。例えば、パターン s^1, s^2, s^3 全てで (+) 状態のニューロンを (+++) 型、パターン s^1, s^2 で (+) 状態、 s^3 で (-) 状態となるニューロンを (++-) 型というように分類することにする。この分類法では、3つのパターンを記憶する場合は、 $2^3 = 8$ 種類のニューロン群に分類される (図7.11)。この8つのニューロン群に $\alpha = 1$ から 8 までの番号をふり、 α 番目のグループでの記憶パターン s^p のとる値を s_α^p 、ニューロンの内部状態を u_α 、出力を x_α と表記する。また、グループ α のニューロンの数を n_α とする。この各 n_α の値がパターン間の相関構造を決定する。

親パターンと子パターンの相関が r の場合、子パターンどうしの相関は $R = r^2$ である。また、各グループに属するニューロン数 n_α は、 $q = \frac{1}{2}(1+r)$ とすると

$$n_{(+++)} = q^3 \quad (7.15)$$

$$n_{(++-)} = q^2(1-q) \quad (7.16)$$

$$n_{(+--)} = q^2(1-q) \quad (7.17)$$

$$n_{(-++)} = q^2(1-q) \quad (7.18)$$

$$n_{(+--)} = q(1-q)^2 \quad (7.19)$$

$$n_{(-+-)} = q(1-q)^2 \quad (7.20)$$

$$n_{(--+)} = q(1-q)^2 \quad (7.21)$$

$$n_{(---)} = (1-q)^3 \quad (7.22)$$

となる。

連続時間モデルの場合は、更に、同じグループ α に属するニューロンも、その初期値によって二分される。初期値が (+) 状態のニューロンの内部状態を u_α^+ 、出力を x_α^+ と表現し、初期値が (-) 状態のニューロンの内部状態を u_α^- 、出力を x_α^- と表現することにする。また、グループ α において、初期値として (+) 状態をとるニューロンと (-) 状態をとるニューロンの割合をそれぞれ、 r_α^+, r_α^- ($r_\alpha^+ + r_\alpha^- = 1$) とおく。すると、想起ダイナミクスは、

$$\frac{du_\alpha^\pm}{dt} = -u_\alpha^\pm + \sum_{p=1}^3 s_\alpha^p m_p \quad (7.23)$$

$$m_p = \frac{1}{N} \left(\sum_{\alpha=1}^8 r_\alpha^+ n_\alpha s_\alpha^p x_\alpha^+ + \sum_{\alpha=1}^8 r_\alpha^- n_\alpha s_\alpha^p x_\alpha^- \right) \quad (7.24)$$

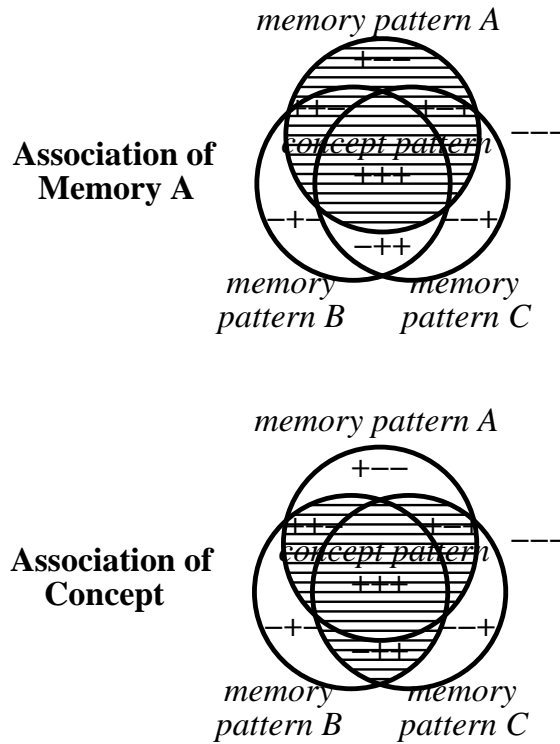


図 7.11: 有限系解析におけるニューロンの分類。

$$x_{\alpha}^{\pm} = f(u_{\alpha}^{\pm}) \tag{7.25}$$

で与えられる (復号同順)。

この連立微分方程式で、 s^1 に近いパターンからの想起の挙動を、単調ニューロンの場合と非単調ニューロンの場合について計算した結果を図 7.12 に示す。この例は、それぞれ、記憶パターンとその一つ上位の概念パターンとが安定化された場合であるという意味で、三段の相関の例で記憶パターンと第二概念パターンを安定化した場合と対応させて考えることができる。図 7.12 から、有限系の微分方程式によって、概念パターン想起時には記憶パターンに近いパターンが一度想起されるなど、シミュレーションに近い挙動が現れることが分かる。

次に、この想起過程の間に、各グループのニューロンの内部状態がどのように変化するかを調べる。図 7.13 に、有限系解析で得られる単調ニューロン (図 7.13 (a))、非単調ニューロン (図 7.13 (b)) を用いた想起時のニューロンの内部状態の変化を示す。ここでは、想起で特徴的な振舞いをするニューロン群 (+ + +) 型ニューロンと (- + +) 型ニューロンを取り出して図示している。このように、単調ニューロンを用いた場合は、グループ (+ + +) のニューロンは常に (+) を出力する。一方、(- + +) ニューロンは (-) 状態

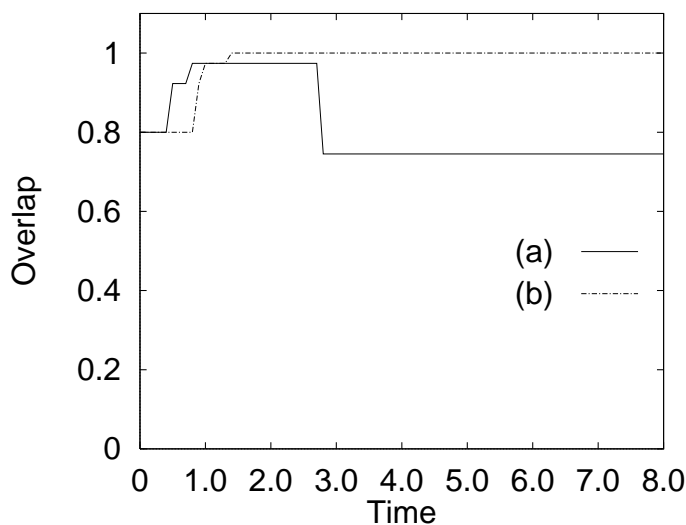
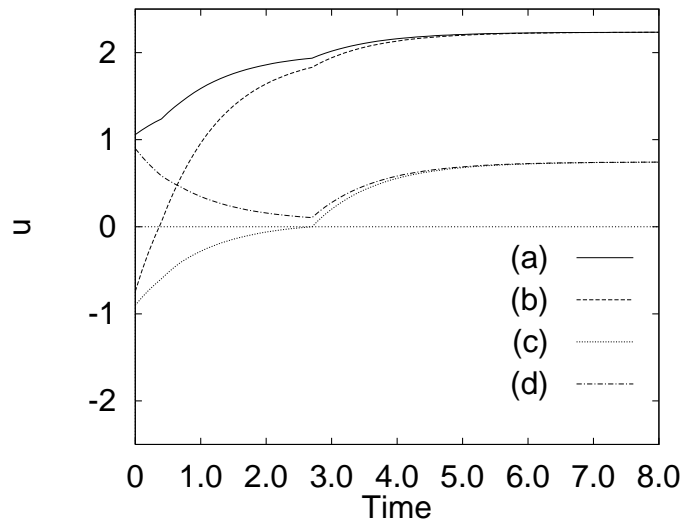


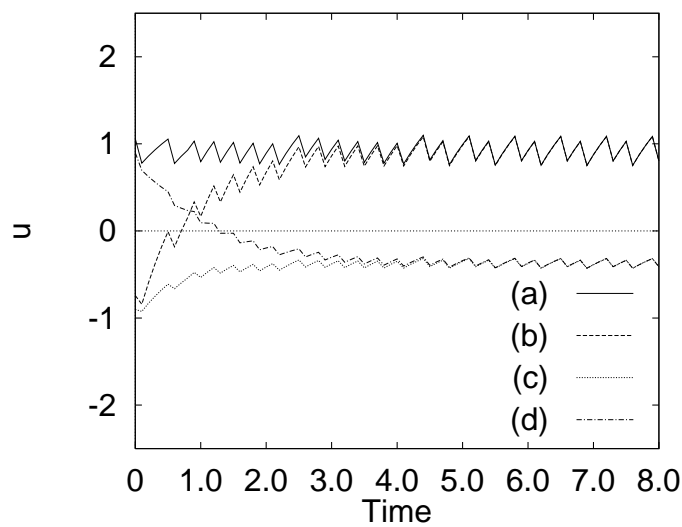
図 7.12: 有限系の微分方程式で求めたオーバーラップの変化 ($R = 0.49, m_1(0) = 0.8$): (a) 単調ニューロン; (b) 非単調ニューロン ($h = 1.0$)。

に留まることができず、概念パターンへと引き込まれてしまう。しかし、非単調ニューロンを使う場合は、(+++)型ニューロンが時々(-)の出力をする一方、(-++)ニューロンの内部状態の変化は抑制され、記憶パターン s^1 が安定化されている。

以上の結果を解釈すると次のような説明が得られる。図7.11から分かる通り、記憶と概念の想起の差は(+--)型ニューロンと(-++)型ニューロンの発火パターンによって決定される。相関が十分大きい場合、(+++)型ニューロンは最も多く、そのため他のニューロンへの影響度(結合荷重の和)も大きい。相関学習では、(+++)ニューロンは(+--)ニューロンと負の結合、(-++)ニューロンとは正の結合を持つ。よって、(+++)ニューロンが発火していると、(-++)ニューロンは発火しやすく、(+--)ニューロンは逆に発火を抑制される。そのため、単調ニューロンを用いた通常の想起では概念パターンが想起されることになる。ところが、非単調ニューロンを用いた場合、内部状態の大きな(+++)ニューロンは時々逆向きの出力を出す。それによって、(+++)ニューロンの影響を弱められ、結果として記憶パターンが安定化されることになるのである。



(a) 単調ニューロンを用いた想起



(b) 非単調ニューロンを用いた想起 ($h = 1.0$)

図 7.13: 想起過程の内部状態の変化 ($m_1(0) = 0.8, R = 0.49$): (a) 初期状態 (+) の (+++) 型ニューロン; (b) 初期状態 (-) の (+++) 型ニューロン; (c) 初期状態 (+) の (-++) 型ニューロン; (d) 初期状態 (-) の (-++) 型ニューロン。

7.3 連続時間相互連想記憶

この節においては、連続時間ダイナミクスを持つ相互相関連想記憶について議論する。相互相関連想記憶の場合は、自己相関連想記憶と異なり、安定点ではなくリミットサイクルが記憶に対応する。離散時間モデルの場合、系の時間発展が過去の内部状態を引きずらないため、オーバーラップのほとんどないパターン間の相互相関を記憶しても連想記憶が実現できた。しかし、同じ荷重行列の構成方法で連続時間ダイナミクスにおけるリミットサイクルを生成することは不可能である。そのため、連続時間力学系では徐々に変化する時系列パターンを相関学習で記憶することになる。以下の議論においては、まず連続時間相互相関連想記憶の定式化を行い、続いて連続時間相互相関連想記憶におけるニューロウインドウ法について述べる。

7.3.1 連続時間相互相関連想記憶

離散時間モデルの場合は、第4章で述べたように隣り合う時間のパターンが大きく異なるような時系列を相関学習で記憶してもその想起が可能である。しかし、連続時間モデルの場合はシステムの内部状態が徐々に変化するので、隣り合う時間のパターンはある程度近いパターンになっている必要がある。そこで、上記のようなランダムパターンの系列を記憶するには、隣り合う時間のパターンに関して、

$$E[\mathbf{s}^{(p)(\tau+\frac{\sigma}{S})} \mathbf{s}^{(p)(\tau)}] = \frac{\sigma}{S} \quad (7.26)$$

$$E[\mathbf{s}^{(p)(\tau+\frac{\sigma}{S})} \mathbf{s}^{(p)((\tau+1)\bmod Q)}] = \frac{S-\sigma}{S} \quad (7.27)$$

を満たす $(S-1)$ 個のパターン $\mathbf{s}^{(p)(\tau+\frac{\sigma}{S})}$ ($\sigma = 1, 2, \dots, S-1$) を生成し、これらの中間パターンを間に挟んだ系列

$$\begin{aligned} \mathbf{s}^{(p)(0)} &\rightarrow \mathbf{s}^{(p)(\frac{1}{S})} \rightarrow \mathbf{s}^{(p)(\frac{2}{S})} \rightarrow \dots \rightarrow \mathbf{s}^{(p)(\frac{S-1}{S})} \rightarrow \mathbf{s}^{(p)(1)} \\ &\rightarrow \mathbf{s}^{(p)(1+\frac{1}{S})} \rightarrow \dots \rightarrow \mathbf{s}^{(p)(Q-1+\frac{S-1}{S})} \rightarrow \mathbf{s}^{(p)(0)} \end{aligned}$$

を相関学習で記憶するようにする。すなわち、結合荷重を

$$w_{ij} = \frac{1}{N} \sum_p \left(\sum_{\sigma=0}^{SQ-2} s_i^{(p)(\frac{\sigma+1}{S})} s_j^{(p)(\frac{\sigma}{S})} + s_i^{(p)(0)} s_j^{(p)(\frac{SQ-1}{S})} \right) \quad (7.28)$$

($i \neq j$) のように設定する。

しかし、結合荷重をこのように設定して、ダイナミクスとして通常の単調ニューロンを用いると、記憶した軌道を想起することはできない。その想起を可能にするために非単調ニューロンを用いる方法が提案されている [30]。そのメカニズムについて、次の小節でニューロウインドウ法による想起メカニズムと合わせて論じることにする。

7.3.2 連続時間相互相関連想記憶の幾何

前節 7.2 で述べたように、連続時間型自己相関連想記憶においては、非単調ニューロンの窓の大きさをニューロウインドウ法で制御することで、記憶パターンとその平均パターン（概念パターン）の想起を選択的に行うことが可能である。ここでは、そのような選択的想起を前小節で定義した連続時間の相互想起型連想記憶において実現する方法を検討する。

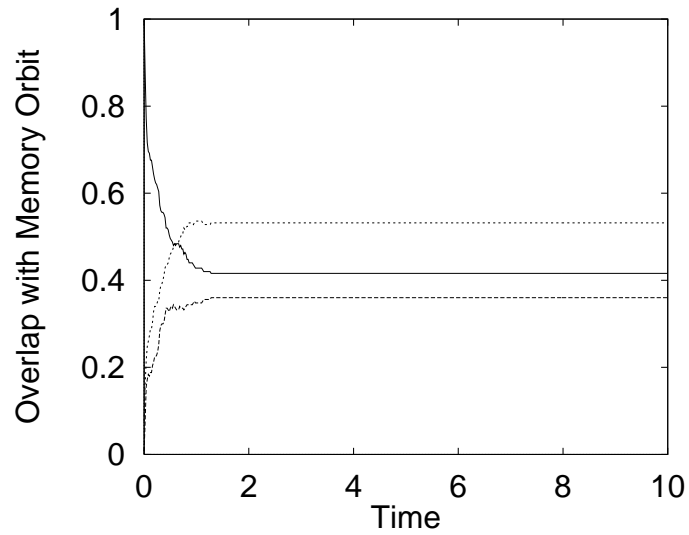
連続時間モデルの想起においては、非単調ニューロンの窓の大きさを決定するパラメータ h 以外に、ニューロンの時定数 τ が可変パラメータとして存在する。しかし、 τ を変化させても、系の変化のスピードが変わるだけなので、以下では $\tau = 1$ としてパラメータ h を変えたときの想起の様子を調べることにする。ここでは簡単のため、ニューロン数 $N = 500$ 、相関の段数 $L = 2$ 、クラスターの相関 $R = 0.64$ 、中間パターン数 $S = 10$ 、パターン数 $27 (P_1 = 3, P_2 = 3)$ の場合をとりあげて考えることにする。ここで取り上げる例は相関の段数が 2 であるから、平均軌道は一階層だけ定義される。このような相関の階層構造をもつ記憶系列に対し、非単調ニューロンのパラメータ h を振ってその想起過程を数値シミュレーションした結果を図 7.14、図 7.15、図 7.16 に示す。

まず、 h が非常に大きい場合 ($h = 50$) であるが、この場合は軌道の想起に完全に失敗し、ある状態に収束して状態遷移が止まってしまう（図 7.14）。次に、 h を多少小さめにした場合 ($h = 10$) は平均軌道に近い軌道が想起される（図 7.15）。更に h を小さくすると ($h = 2.0$)、今度は記憶軌道に近い軌道が想起される（図 7.16）。このように、非単調ニューロンのパラメータ h を変化させることで、平均軌道と記憶軌道を選択的に想起することができることが数値実験で確認された。

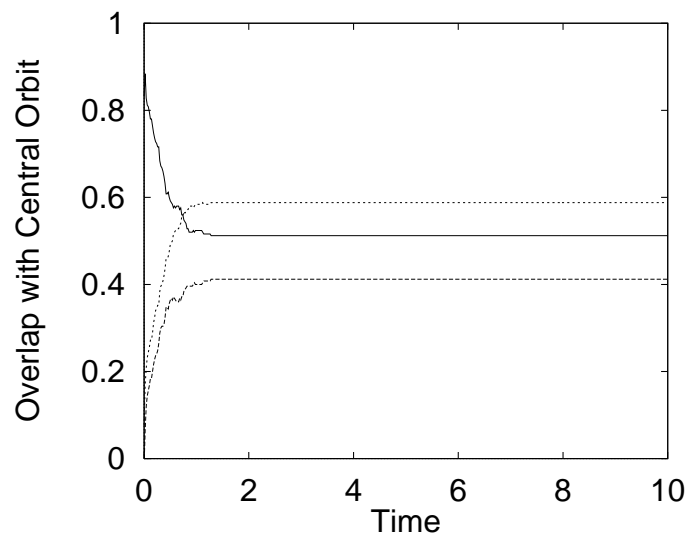
さて、以上の議論で確認された連続時間相互相関連想記憶の性質に対し、幾何的な視点から考察を加えることにする。自己相関連想記憶の場合と異なり、相互相関連想記憶の場合は、連続時間モデルと離散時間モデルで記憶行列自体が異なる。そこで、記憶行列の固有空間解析から始めることにする。

W は相互相関行列なので非対称成分をもつ。よって、行列 W は対称成分 $[W]_s$ と非対称成分 $[W]_a$ に分解できる。ここで、 W は相互相関行列といえども、近いパターン間の相互相関をとるので、生成される行列は自己相関行列に近く、対称成分が大きくなる。例えば、上の数値実験で用いた $S = 10$ の例では、時系列中で連続するパターン間の相関の期待値は $\frac{S-1}{S} = 0.9$ となり、95% は同じ出力を出すことになる。よって、対称成分と非対称成分の割合は 19 : 1 となる。実際、上記の数値実験で用いた W の固有値分布を調べた結果を図 7.17 に示す。この図のように、どの固有値も位相はほぼ 0 か π になっており、行列は対称に近いことが見て取れる。

次に対称成分のうちわけを考える。対称成分はパターンの自己相関に対応する部分である。この対称成分の固有値の分布を調べると図 7.18 のようになる。このように、固有値の分布は段差を生じている。これは自己相関連想記憶の場合と同じである。次に、記憶軌道中のパターン、平均軌道中のパターン、及び偽記憶パターン（ h が大きく状態遷移が収束する時の収束先のパターン）を固有ベクトルに分解した時の各成分の大きさを図 7.19(a),(b),(c) に示す。記憶軌道中のパターンと平均軌道中のパターンは相関がある

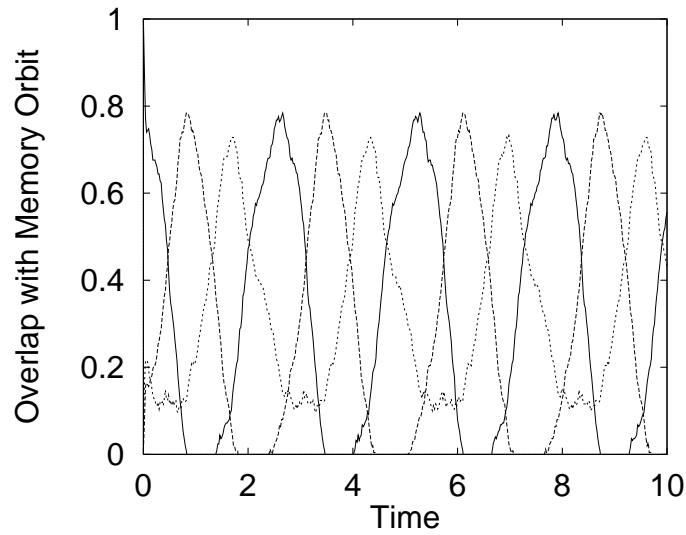


(a) 記憶軌道とのオーバーラップ

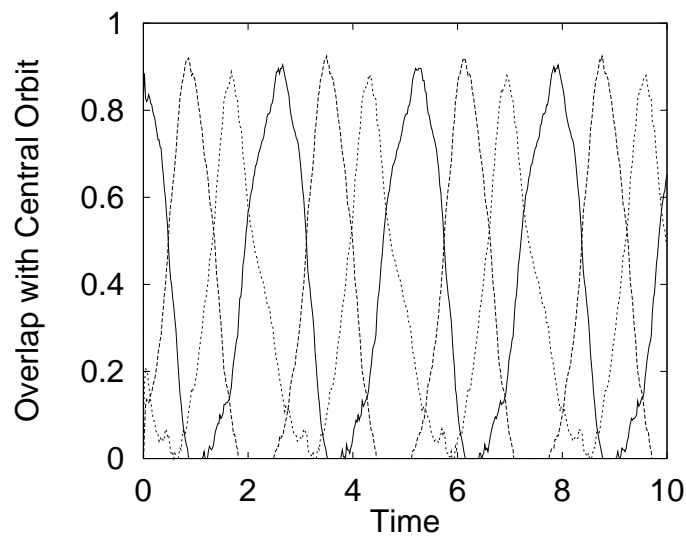


(b) 平均軌道とのオーバーラップ

図 7.14: $\hbar = 50$ に設定した時の系の時間発展。図中の3つの線はそれぞれ軌道中の1つのパターンとのオーバーラップの変化を示している。この場合は系の状態遷移は途中で停止してしまい、どの軌道も想起できない。

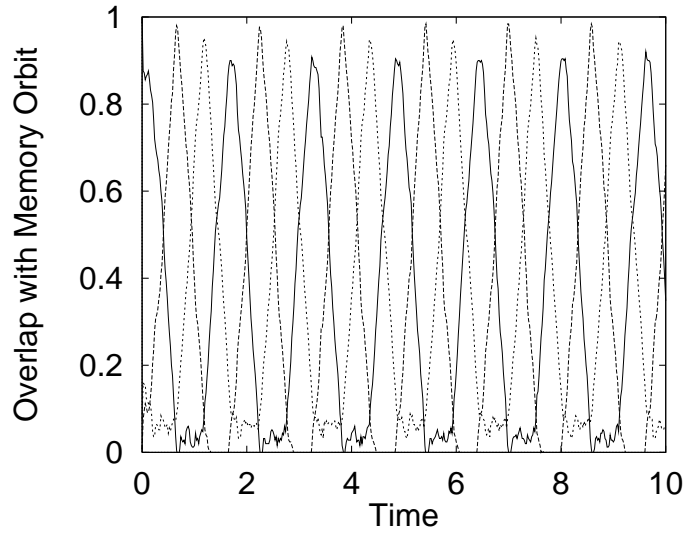


(a) 記憶軌道とのオーバーラップ

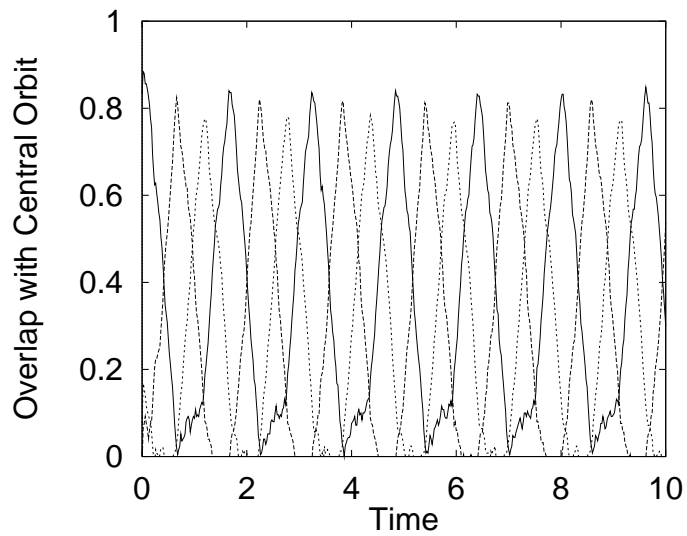


(b) 平均軌道とのオーバーラップ

図 7.15: $h = 10$ に設定した時の系の時間発展。図中の3つの線はそれぞれ軌道中の1つのパターンとのオーバーラップの変化を示している。ここでは平均軌道に近い軌道の想起が実現されている。



(a) 記憶軌道とのオーバーラップ



(b) 平均軌道とのオーバーラップ

図 7.16: $h = 2.0$ に設定した時の系の時間発展。図中の3つの線はそれぞれ軌道中のパターンの1つのオーバーラップの変化を示している。ここでは記憶軌道に近い軌道の想起が実現されている。

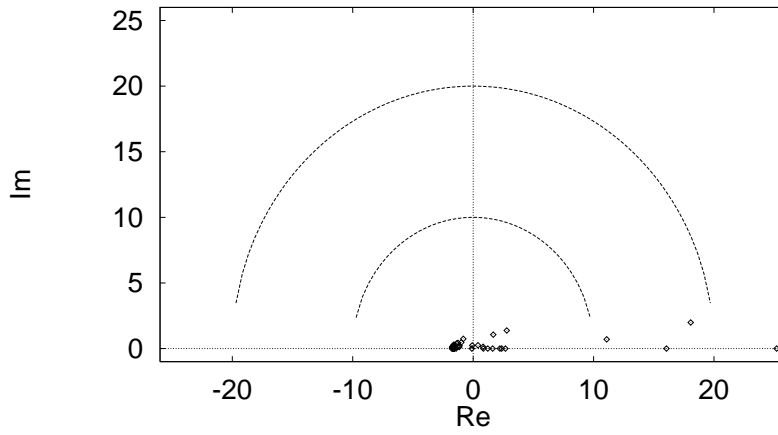


図 7.17: 記憶行列 W の固有値の分布 ($N = 500, R = 0.64, L = 2, P_1 = 3, P_2 = 3, Q = 3, S = 10$)。

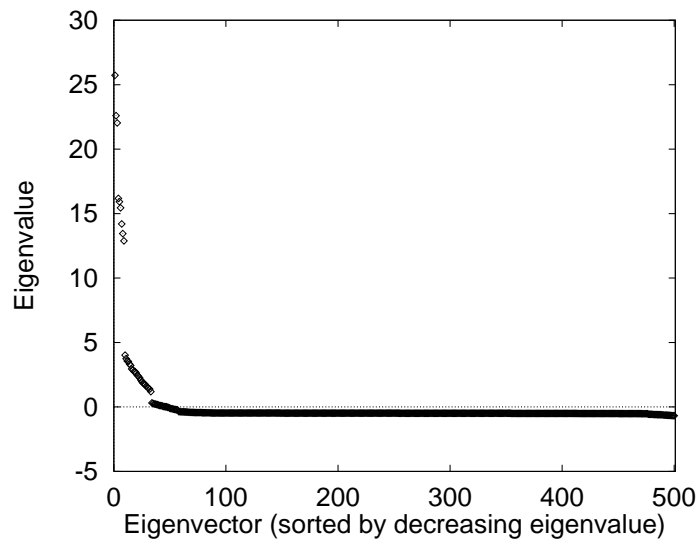
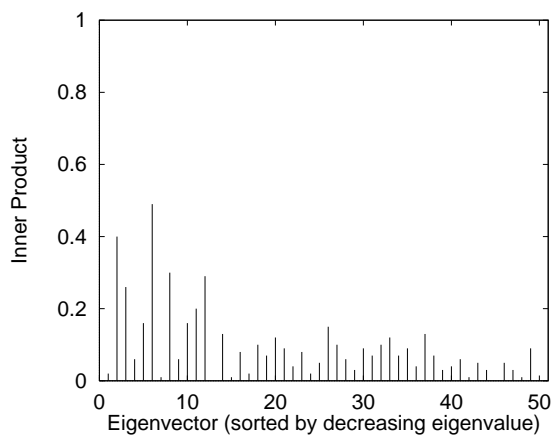
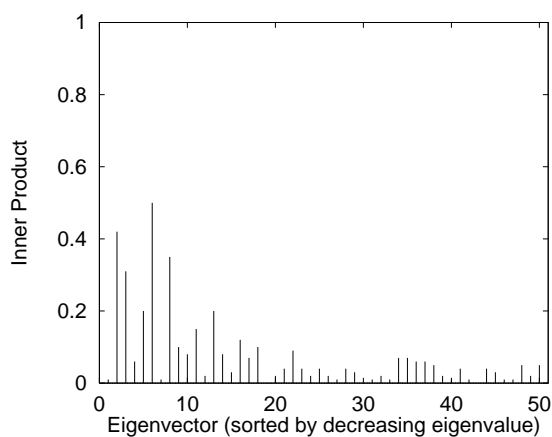


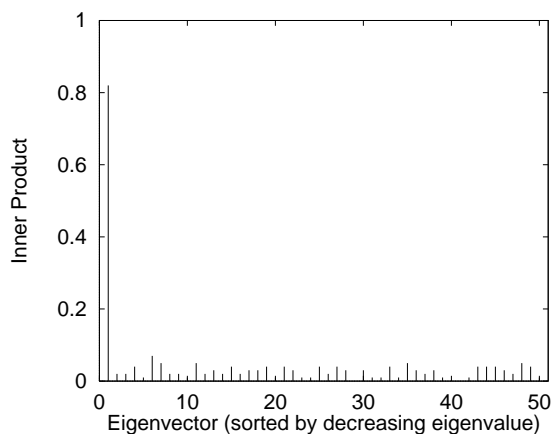
図 7.18: 記憶行列 W の対称成分 $[W]_s$ の固有値の分布 ($N = 500, R = 0.64, L = 2, P_1 = 3, P_2 = 3, Q = 3, S = 10$)。



(a) 記憶軌道中のパターンの固有空間における表現



(b) 平均軌道中のパターンの固有空間における表現



(c) 偽記憶パターンの固有空間における表現

図 7.19: 記憶軌道中のパターン、平均軌道中のパターン、及び偽記憶パターン (h が大きく状態遷移が収束する時の収束先のパターン) と $[W]_s$ の固有ベクトルの内積。固有値が大きい上位 100 個の固有ベクトルとの内積を示している。

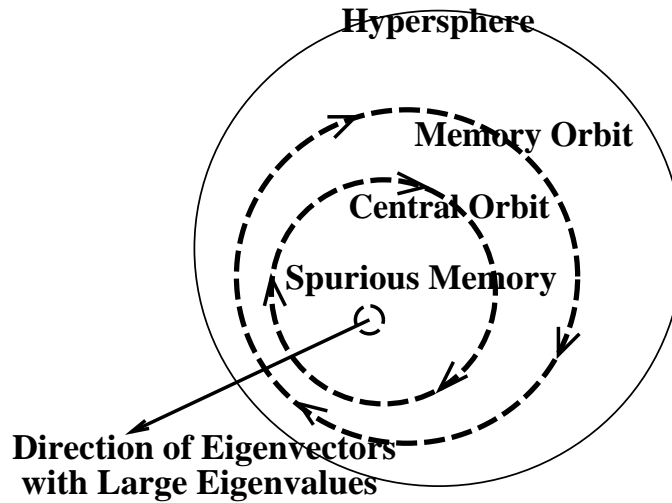


図 7.20: 軌道の幾何的模式図。偽記憶は最大固有値を持つ固有ベクトルの方向に存在し、平均軌道は大きい固有値をもつ空間で回転、記憶軌道はそれより小さめの固有値をもつ空間で回転している。ニューロウィンドウ法では h を小さくすることで固有値の大きい空間から順に不安定化できるので、平均軌道や記憶軌道を選択的に安定化することができる。

ので似た成分を持つが、平均軌道中のパターンの方が大きい固有値を持つ固有ベクトル成分が大きくなっている。一方、偽記憶パターンは最大固有値を持つ固有ベクトル成分が突出して大きくなっている。これも自己相関連想記憶の場合と同様の現象である。

以上のことから、これらの軌道がなす幾何構造は図 7.20 のようになっていると言える。非単調ニューロンは内部状態が h 以上になるとそのニューロンの出力の影響を切ってしまう。内部状態が大きい状態は、大きい固有値を持つ固有ベクトルで張られる空間に存在するから、非単調ニューロンは大きい固有値を持つ固有ベクトルへの流れを止める効果を持つ。よって、 h が大きいと流れの塞き止めが起きず、最大固有値を持つ固有ベクトルの方向へ状態が遷移していく。 h を小さくしていくと、より小さな固有値で張られる空間が安定化していくため、平均軌道、更には記憶軌道が安定化していくことになる。

もちろん、相互想起型の場合は、単に行列の対称成分の大きい固有値をもつ固有ベクトル方向に進むだけではなく、回転要素も入っている。よって、 h を小さくすることにより回転要素が不安定化される恐れもある。しかし、上で述べたように非対称成分は対称成分に比べるとごく小さく、平均軌道や記憶軌道を安定化するレベルの h の値では回転を不安定化するような影響は無い。

また、数値実験では、同じ時定数 τ を使っても h の値が小さいときほどパターンの変化が速くなるという現象が見られる。これは、 h が小さいときは内部状態の項 u が大きくなるため、入力項の効果が早く現れることによると説明できる。

7.4 ランダムネットワークの連続時間ダイナミクス

前節で述べた相互相関連想記憶の場合は、非対称ネットワークといっても、非対称成分が比較的小さい場合を考えたので、対称成分だけに着目して議論を行うことができた。しかし、一般には非対称成分は大きく、非対称成分も含めた議論が必要になる。そこで、この節ではランダムネットワークを取り上げ、その連続時間モデルの挙動を解析する。

ランダムネットワークの固有空間の性質、及びその挙動は既に離散時間モデルの時に述べた。ここでは、離散時間モデルと連続時間モデルの差を同じランダム行列を使った場合で比較することにする。図 7.21、表 7.1 に示す固有値を持つランダムネットワークにおいて、離散時間、及び連続時間ダイナミクスで系を遷移させた場合の固有ベクトル成分の変化を調べた結果を、それぞれ図 7.22、図 7.23 に示す。この結果から分かるように、離散時間モデルは固有値の絶対値が大きい空間が支配的であるが、連続時間モデルでは固有値の実部が大きい空間が支配的であり、絶対値が大きくても実部が小さい値を取る空間の成分は低く抑えられている。

この現象のメカニズムは以下の通り説明できる。今まで述べてきた通り、ニューロダイナミクスでは非線形変換はローカルであり、線形変換がドミナントに働く。よって、ニューロンが線形素子であると仮定してみよう。すると、そのダイナミクスは

$$\dot{u}_i = -u_i + \sum_j w_{ij} u_j \quad (7.29)$$

で与えられ、その解は

$$\mathbf{u} = \sum_{\mu} e^{\lambda_{\mu} t} \mathbf{f}^{\mu} \quad (7.30)$$

で与えられる。 λ_{μ} を実部、虚部に分け、

$$\mathbf{u} = \sum_{\mu} e^{Re[\lambda_{\mu}]t} e^{Im[\lambda_{\mu}]t} \mathbf{f}^{\mu} \quad (7.31)$$

と書き直すと、 $e^{Im[\lambda_{\mu}]t}$ は振動を生み、 $e^{Re[\lambda_{\mu}]t}$ は増幅、又は減衰を促す。よって、 $Re[\lambda_{\mu}]$ の大きさが固有ベクトル \mathbf{f}^{μ} の成分の増減を決定する。但し、単純な線形系では $Re[\lambda_{\mu}]$ の正負で増減が分類できるが、ニューロでは非線形素子によるベクトルの長さの規格化操作が加わるので、全固有値を並べたときの $Re[\lambda_{\mu}]$ の比が各固有ベクトル成分の増減を決定する。その結果、連続時間ニューロダイナミクスでは、固有ベクトル群の中で $Re[\lambda_{\mu}]$ の大きいものが支配的になるのである。

以上のことは、離散時間と対比して以下のように説明することも可能である。離散時間モデルの場合、非線形効果を無視すると、状態遷移は

$$\mathbf{u}(t+1) = W\mathbf{u}(t) \quad (7.32)$$

であり、 W そのものが状態遷移の作用素となっていた。ところが、連続時間モデルでは

$$\mathbf{u}(t+dt) = \left((1-dt)I + dtW \right) \mathbf{u}(t) \quad (7.33)$$

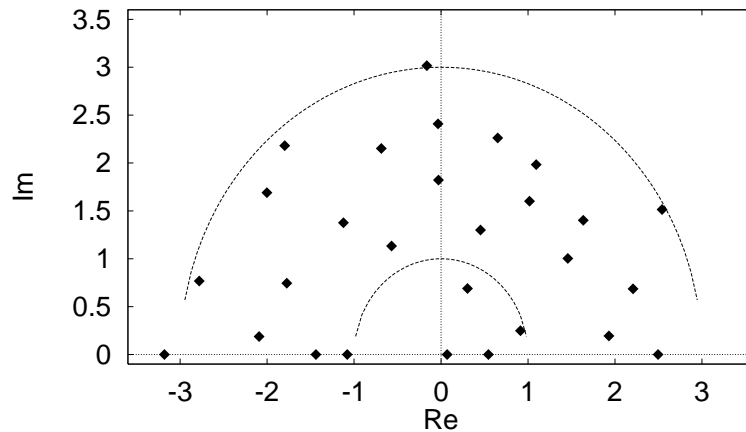


図 7.21: ランダムネットワークの固有値。ここでは各成分が $[-1,1]$ の一様分布に従う 50 次元ネットワークの全固有値を示している。

固有値番号	1	2	3	7	10
絶対値	3.18	3.02	2.96	2.50	2.31
実部	-3.18	-0.16	2.54	2.50	2.21

表 7.1: 上の実験で用いたランダム行列によるダイナミクスで注目した固有ベクトルの固有値。固有値番号は絶対値の大きい順。

であり、 $(1 - dt)I + dtW$ が作用素となっている。 dt が小さければ、 W の大きい固有値をもつ固有ベクトル方向は $(1 - dt)I + dtW$ の絶対値の大きな固有値をもつ固有ベクトル方向に対応する。このことは、本節で述べた離散時間モデルと連続時間モデルの性質の違いの根本をよく表している。

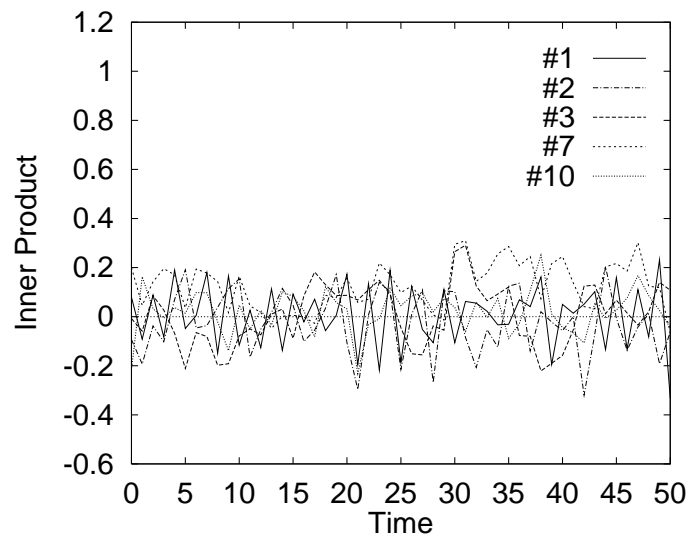


図 7.22: ランダムネットワークにおける離散時間ダイナミクスでの固有ベクトル成分の変化。図中#は固有値番号を表す。

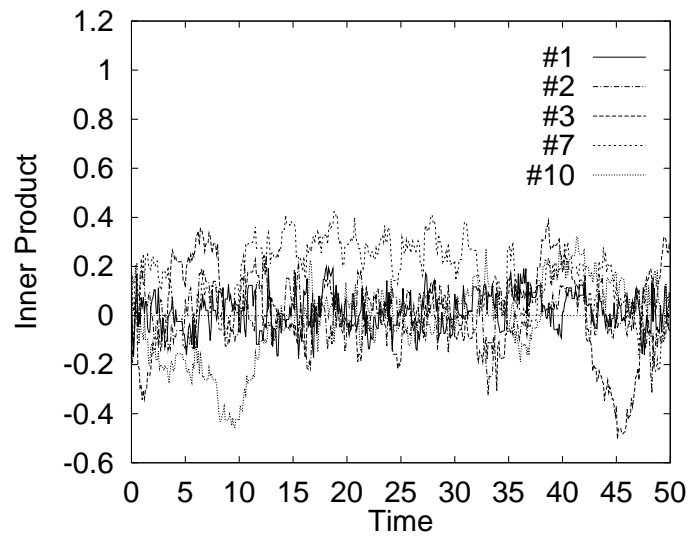


図 7.23: ランダムネットワークにおける連続時間ダイナミクスでの固有ベクトル成分の変化。

7.5 まとめ

本章では、連続時間ダイナミクスを持つニューラルネットワークの性質について、幾何的な議論を中心に、離散時間モデルと比較しながら解析を行った。

荷重行列が対称な場合においては、連続時間モデルでは系は固有値が正の大きな値を持つ固有ベクトル方向へと状態遷移する。その流れは、非単調ニューロンを用いることで抑制されるので、連続時間自己相関連想記憶においても、非単調ニューロンの閾値を調節することで不安定領域を制御し、ニューロウインドウ法による選択的想起が実現できることを示した。また、系が連続的に変化することから現れる引き込み領域の増大や下位概念の一時的想起現象のメカニズムを説明した。

相互相関連想記憶の場合、連続時間モデルでは系の状態遷移が連続的であるので、近いパターン間の相関をとるように記憶法自体を変える必要がある。そのため、荷重行列の固有空間の性質自体に変化が生じる。近いパターンの相互相関を取ると、結果として得られる荷重行列はほぼ対称となるので、その対称成分を中心に議論を行うことができる。よって、荷重行列が対称な系と全く同じ形で非単調ニューロンを用いることで、ニューロウインドウ法による選択的想起が実現できることができた。

荷重行列の非対称性が強い場合として、ランダムネットワークの連続時間ダイナミクスを調べると、系は実部の大きい固有値をもつ固有ベクトルに支配されることが分かった。

第 8 章： 組み合わせ最適化の幾何とその応用

これまでの議論においては、ニューラルネットワークのダイナミクスとして、主に連想記憶を扱ってきた。その他に、ニューラルネットワークのダイナミクスを利用したシステムで代表的なものとして、組み合わせ最適化問題を解くホップフィールドネットワークが存在する。本章では、その組み合わせ最適化問題を解くネットワークに対して幾何的な視点から解析を加える。

前章の議論では、連続時間ダイナミクスの離散時間ダイナミクスとの相違点について述べた。組み合わせ最適化問題の場合、既存の解探索のダイナミクスとしてはシミュレーテッドアニーリングによる確率的探索 [21]、或は連続時間ダイナミクス [18] が用いられる。本章では、幾何的視点をもとに、並列デジタル計算という条件下で、従来の探索手法より高速に準最適解を得る方法を提案する。

まず、第1節では、従来のニューラルネットワークによる組み合わせ最適化手法をレビューする。第2節では、 $(-1, 1)$ 出力型ニューロンで表現される組み合わせ最適化問題の例として分割問題を取り上げてその幾何構造を解析し、並列デジタル計算による準最適解の高速探索法を提案する。第3節では、 $(0, 1)$ 出力型ニューロンで表現される組み合わせ最適化問題の例として巡回セールスマン問題を取り上げ、その高速解法を幾何的視点から検討する。

8.1 ニューロによる組み合わせ最適化

ニューロによる組み合わせ最適化においても、連想記憶と同様 N 個の全ニューロンどうしが互いに重みづけして結合されたネットワーク (ホップフィールドネットワーク) を考える。即ち、第 i ニューロンは x_i を出力し、入力

$$u_i = \sum_{j=1}^N w_{ij} x_j \quad (8.1)$$

を受け取る。ニューロンは、この入力

$$x_i = f(u_i) \quad (8.2)$$

のように更新する。但し、連想記憶では $f(u)$ として主に符号関数を考えたが、組み合わせ最適化のダイナミクスの場合、通常 $(0, 1)$ モデルではヘビサイド関数

$$f(u_i) = \frac{1}{1 + \exp(-\beta u_i)} \quad (8.3)$$

を用い、 $(0, 1)$ モデルではシグモイド関数

$$f(u_i) = \frac{1 - \exp(-\beta u_i)}{1 + \exp(-\beta u_i)} \quad (8.4)$$

を用いる。ここで、 β はアニーリングに使われる温度パラメータ T の逆数である。この節では、簡単のため $(-1, 1)$ モデルを仮定して議論を進めることにする。

結合が対称な場合、すなわち $w_{ij} = w_{ji}$ の場合、シミュレーテッドアニーリングやアナログニューロンを用いた状態更新則(後述)のもとでは、関数

$$E(\mathbf{x}) = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N w_{ij} x_i x_j \quad (8.5)$$

は系のエネルギー関数(リアプノフ関数)になっていることが知られている。ホップフィールドネットワークを用いて組み合わせ最適化問題を解く場合、状態 \mathbf{x} が解の一候補を表すようにし、エネルギー $E(\mathbf{x})$ の最小値が最適解に対応するように荷重行列 $W = [w_{ij}]$ を設計する。

組み合わせ最適化問題は、小さくすべきコストと、解が満たすべき制約条件の二つの要素から成り立っている。解の候補 \mathbf{x} について、そのコスト $C(\mathbf{x})$ と制約条件を破ったときのペナルティ $P(\mathbf{x})$ が \mathbf{x} の二次形式で表現できる場合、荷重行列 W を

$$E(\mathbf{x}; W) = C(\mathbf{x}) + P(\mathbf{x}) \quad (8.6)$$

が満たされるように設計すれば、系のエネルギー関数の減少を使って解の改善を行っていくことができる。

以下では、エネルギー関数を減らすダイナミクスとして、シミュレーテッドアニーリング法 [21] とアナログニューロンを用いたホップフィールドネットワーク [18] を紹介する。

Kirkpatrick のシミュレーテッドアニーリング法では、 N 個のニューロンからランダムに一個ずつニューロンを取りだし、そのニューロンの状態 x_i を確率

$$P(x_i(t+1) = 1) = \phi\left(\sum_{j=1}^N w_{ij} x_j(t)\right) \quad (8.7)$$

$$P(x_i(t+1) = -1) = 1 - \phi\left(\sum_{j=1}^N w_{ij} x_j(t)\right) \quad (8.8)$$

$$\phi(u) = \frac{1}{1 + \exp(-\beta u_i)} \quad (8.9)$$

に従って更新する。しかし、この方法では、問題が並列にインプリメントされているにもかかわらず、シリアルにしか計算が行えない。

Hopfield と Tank のアナログホップフィールドネットワークでは

$$\dot{u}_i = -\frac{1}{\tau} u_i + \sum_{j=1}^N w_{ij} x_j \quad (8.10)$$

$$x_i = f(u), \quad (8.11)$$

$$f(u) = \frac{1 - \exp(-\beta u)}{1 + \exp(-\beta u)} \quad (8.12)$$

という微分方程式を用いる。これは、前章で述べたニューロダイナミクスの連続時間モデルである。よって、並列計算を行うことが可能となる。但し、この形では x が 1 と -1 に離散化されていないので、解自体は

$$y_i = \text{sgn}(x_i), \quad (8.13)$$

$$\text{sgn}(x) = \begin{cases} 1, & \text{if } x \geq 0, \\ -1, & \text{if } x < 0 \end{cases} \quad (8.14)$$

で与えられる状態ベクトル \mathbf{y} で判定する。しかし、この Hopfield と Tank の方法はアナログ計算なので、前章で述べた通り、デジタル計算機で上の微分方程式を解くには、差分方程式

$$u_i(t + \tau\Delta) = (1 - \Delta)u_i(t) + \Delta \sum_{j=1}^N w_{ij}x_j(t) \quad (8.15)$$

に一度変換してから計算する必要がある。この差分方程式で微分方程式の解の振る舞いを近似的に再現するためには差分を細かくとることが要求されるが、そうすると今度は並列計算にしてもある程度の計算時間が必要になってしまう。この問題点を解決する方策として、差分を大きくとることが考えられる。最も極端なケースとしては、 $\Delta = 1, \tau = 1$ とし、

$$x_i(t + 1) = f\left(\sum_{j=1}^N w_{ij}x_j(t)\right) \quad (8.16)$$

で表される単純な離散時間ダイナミクスを用いることが考えられる。しかしながら、差分を大きくとると、ダイナミクスは振動的になり、関数 $E(\mathbf{x})$ は減少しない。つまり、差分が大きい場合は $E(\mathbf{x})$ は系のリアプノフ関数ではなくなってしまう。

次節では、以上で述べた解探索ダイナミクスの問題を解消する方法について、組み合わせ最適化問題の一つである分割問題を例にとりあげて考える。そして、分割問題において、並列デジタル計算の枠組みで準最適解を高速に得る方法を提案する。

8.2 分割問題の幾何とその応用

8.2.1 分割問題

分割問題 (Partition Problem) とは、互いに配線で結ばれている N 個の素子を、箱の外部の配線 (二つの箱の間の配線) ができるだけ少なくなるように、容量が $N/2$ の二つの箱 A, B につめる問題である (図 8.1)。以下では、素子 i と素子 j の配線の数 d_{ij} 本あるとする。この問題では、コストは外部の配線数、制約はそれぞれの箱の中の素子数が箱の容量を越えないことにあたる。

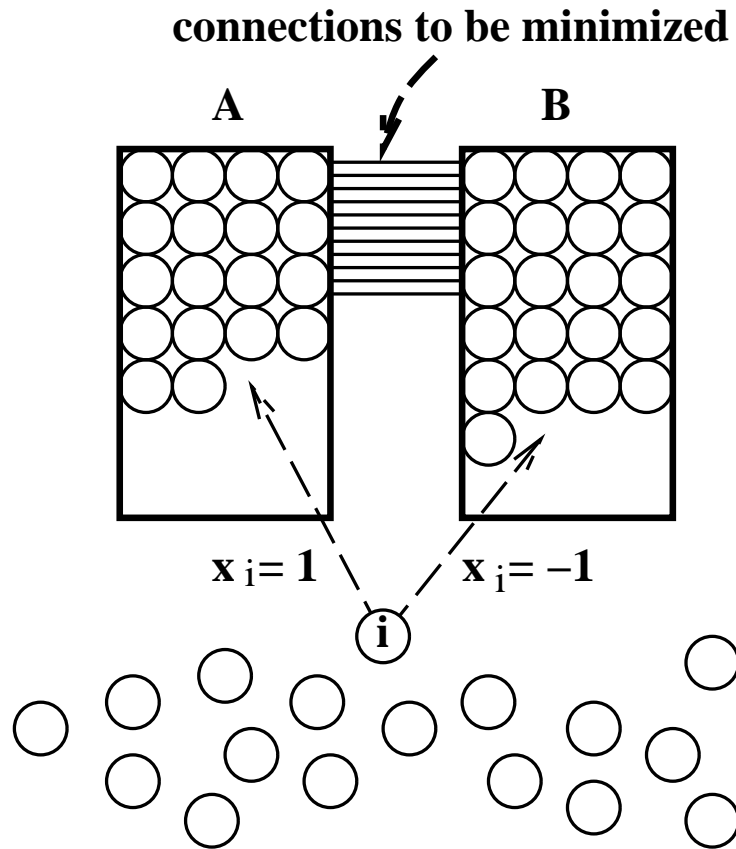


図 8.1: 分割問題。箱の間の配線を最小にするように素子を二つの箱に分けて詰める。

分割問題を解くニューラルネットワークとして、 N 個のニューロンを用意する。これらのニューロンはそれぞれの素子がどちらの箱に入るかを表すことにする。すなわち、 i 番目のニューロンの状態 x_i が 1 なら、 i 番目の素子は箱 A に入り、 $x_i = -1$ なら i 番目の素子は箱 B に入ることにする。

まず、コストの表現であるが、 i 番目の素子と j 番目の素子が異なる箱に入ると d_{ij} 本の配線が箱外に持ち出される。箱外の全配線数 C は

$$C = \frac{1}{8} \sum_{i=1}^N \sum_{j=1}^N d_{ij} (x_i - x_j)^2 \tag{8.17}$$

と表される。 $x_i^2 = 1$ であることに留意し、これを展開すると

$$C = -\frac{1}{4} \sum_{i=1}^N \sum_{j=1}^N d_{ij} x_i x_j + \frac{1}{4} \sum_{i=1}^N \sum_{j=1}^N d_{ij} \tag{8.18}$$

となる。この式の第二項は定数になるから、ネットワークの結合荷重設計の時は第一項だけを考えれば良い。

一方、制約条件を満たさない時のペナルティーであるが、こちらは箱に等分配されたときに最小になる関数を用意すれば良い。そのためには、簡単に、各素子間に均等な斥力を設ければ良い。具体的には、

$$P = \sum_{i=1}^N \sum_{j \neq i}^N h x_i x_j \quad (8.19)$$

という関数をペナルティー項とする。

以上で得られたコスト項とペナルティー項を用い、分割問題を解くネットワークの結合荷重

$$w_{ij} = d_{ij} - h(1 - \delta_{ij}) \quad (8.20)$$

を得ることができる。ここで、 δ_{ij} はクロネッカーのデルタ関数である。

この分割問題も、シミュレーテッドアニーリングやアナログニューロンを用いた連続時間ダイナミクスにより準最適解を得ることができるが、離散時間同期型ダイナミクスや、大きい差分を用いた差分方程式によるダイナミクスを用いると、系が振動的になり準最適解は得られない。次小節の議論では、分割問題を解くホップフィールドネットワークの結合荷重行列について、その固有空間の構造を解析し、離散時間モデルで解探索のダイナミクスが振動的になる原因を調べる。

8.2.2 分割問題の幾何

ここでは、各素子間の配線数は 50 から 150 までの自然数を等確率でとる例を考える。但し、行列を作るときは配線数に 0.01 倍して規格化し、 d_{ij} の平均が 1 になるようにしている。

この分割問題を解くネットワークの荷重行列

$$w_{ij} = d_{ij} - h(1 - \delta_{ij}) \quad (8.21)$$

の固有値の分布を図 8.2 に示す。ここで、素子数 N は 100 であり、 $h = 1.5$ としている。

この図から分かるように、固有値の中に一つだけ絶対値が他の固有値よりも遥かに大きい負の固有値が存在している。この固有値を持つ固有空間は、制約条件のペナルティーを表す行列

$$w_{ij}^p = -h \quad (8.22)$$

によって生成される。この行列のランクは 1 であり、全ての成分が負であるので、この行列は絶対値の大きい負の固有値を持つ 1 次元空間を形成する。これが行列 W の絶対値の大きい負の固有値を作り出している。すなわち、この空間は制約条件を満たさない状態の集合を張る空間であるということができる。

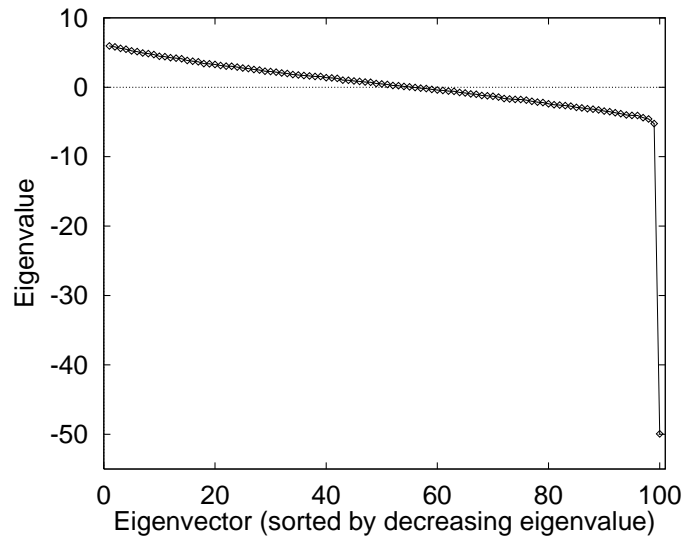


図 8.2: 固有値の分布 ($N = 100$)。固有値の絶対値は最小固有値が飛び抜けて大きい。この固有値を持つ固有ベクトルは制約条件を満たさない空間に対応する。

離散時間非同期型ダイナミクスや連続時間ダイナミクスでは、 $-\mathbf{x}^T W \mathbf{x}$ がリアプノフ関数になっているので、行列 W の大きい固有値を持つ固有ベクトル成分が増えていくことになる。実際、得られる準最適解の状態ベクトルと固有ベクトルの内積を調べてみると、系の状態ベクトルは大きい固有値を持つ固有ベクトル群に張られており、小さい固有値を持つ固有ベクトルとは直交に近い関係にあることが分かる(図 8.3)。

ところが、離散時間同期型ダイナミクスでは、上述の幾何構造の視点から明らかにしたように、大きい固有値をもつ固有ベクトル方向ではなく、絶対値が大きい固有値をもつ固有ベクトル方向へと状態が遷移する。したがって、絶対値の大きい負の固有値が存在する場合、その固有値を持つ固有ベクトル方向へと状態は流れていく。よって、 $-\mathbf{x}^T W \mathbf{x}$ は系のリアプノフ関数にはならず、ダイナミクスを進めることで解の性質改善を行うことができない。実際、離散時間ダイナミクスを用いた場合の状態ベクトルと固有ベクトルの内積を調べてみると、系の状態ベクトルは最小固有値を持つ固有ベクトル成分を大きく持っていることが分かる(図 8.4)。これが離散時間同期型ダイナミクスで解探索が振動的になる原因である。

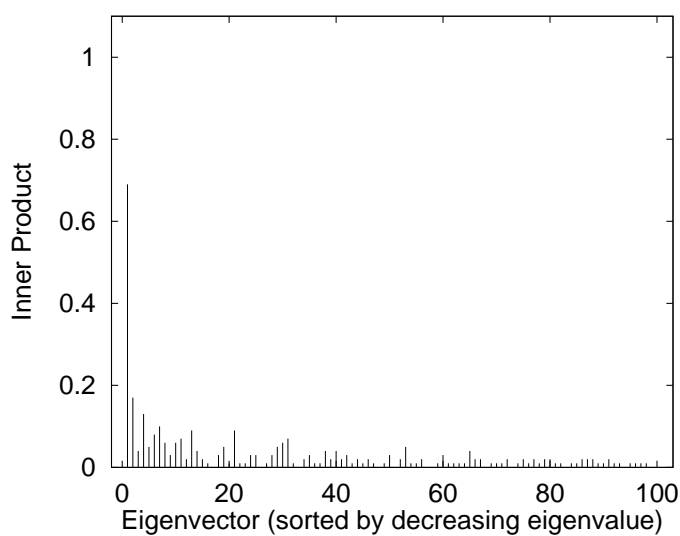


図 8.3: 準最適解の状態ベクトルと結合荷重行列の固有ベクトルの内積 ($N = 100$)。状態ベクトルは大きい固有値をもつ固有ベクトル成分を中心に持つ。

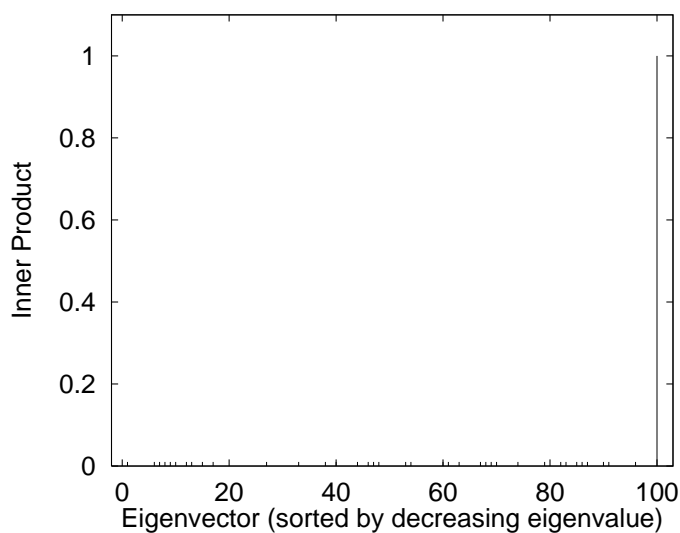


図 8.4: 離散時間同期型ダイナミクスで生起する振動状態の状態ベクトルと結合荷重行列の固有ベクトルの内積 ($N = 100$)。状態ベクトルは最小固有値(絶対値最大の固有値)をもつ固有ベクトル成分がほとんどを占める。

8.2.3 並列同期計算による解の高速探索法

この章では、離散時間同期型ダイナミクスを用いて分割問題の準最適解を高速に得る方法を提案する。

前章で述べた通り、式(8.20)で与えられる標準的な荷重行列 W を使う限りにおいては離散時間同期型ダイナミクスでは解の探索を行えない。つまり、このダイナミクスを使うためには、行列自体に何らかの改良を加えることが必要となる。行列の改良の方針としては次の二点が重要になる。

(条件1) ダイナミクスの収束先の状態ベクトルが W の正の大きな固有値をもつ固有ベクトル成分を大きく持つようにすること。

(条件2) ダイナミクスの収束先の状態ベクトルが W の負で絶対値の大きい固有値をもつ固有ベクトル成分をほとんど持たないようにすること。

この二つの条件を満たすことができれば系のエネルギーは小さくなり、性質の良い解を得ることができる。

これらの条件を満たし、分割問題の準最適解の離散時間同期型探索を可能にする行列として

$$v_{ij} = w_{ij} - \lambda_{\min} e_i^{(\min)} e_j^{(\min)} \quad (8.23)$$

を考える。ここで、 λ_{\min} は最小固有値、 $e^{(\min)}$ はその固有ベクトルである。今、全ての固有値と固有ベクトルに固有値の大きい順に 1 から N まで番号をふることにすると、 $\lambda_{\min} = \lambda_N, e^{(\min)} = e^{(N)}$ と表現できる。

この行列の変更を行うには最小固有値と固有ベクトルを計算する必要がある。最小固有値とその固有ベクトルは時間発展方程式

$$e(t+1) = C_n W e(t) \quad (8.24)$$

をランダムなベクトル $e(0)$ から繰り返し、

$$\lambda e(t+1) \simeq W e(t) \quad (8.25)$$

の状態に収束したときの λ と $e(t)$ として得られる。ここで、 C_n は $|e(t+1)| = 1$ に規格化する定数である。このダイナミクスは、 $|\lambda_N|/|\lambda_i|$ がどの $i (\neq N)$ についても 1 より遥かに大きければ、収束までに要する時間は短い。よって、行列 $V = [v_{ij}]$ は並列計算機があれば高速に得ることができる。

ここで導入した行列 V の固有空間の構造は、行列 W と同じ固有ベクトルをもち、 W の最小固有値をもつ固有ベクトルの固有値だけが 0 に入れ替わった形になっている。ここで得られた行列 V は次に挙げる性質を満たす。

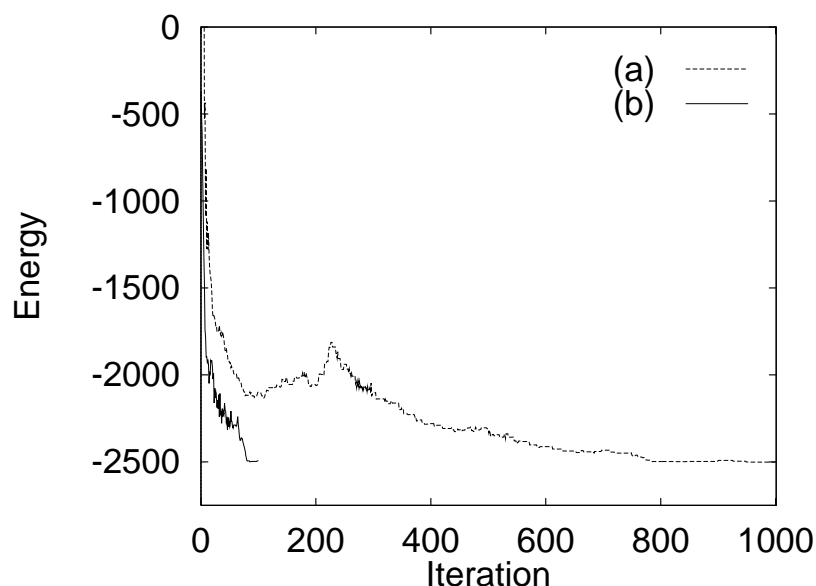


図 8.5: 差分方程式モデル (a) と行列改良型離散同期モデル (b) のエネルギー減少過程 ($N = 500$)。温度パラメータ $T = \beta^{-1}$ については、どちらのモデルにおいても適切なアニーリングを行っている。差分方程式モデルでは $\tau = 1$ とし、 Δ は解探索が最速になる $\Delta = 0.08$ を採用している。この比較の結果から分かるように、行列改良型離散同期モデルは準最適解に高速に収束する。

- (性質1) V の正の大きな固有値が V の負の固有値よりその絶対値が大きいため、離散時間同期型ダイナミクスでも状態ベクトルを V の正の大きな固有値 (= W の正の大きな固有値) をもつ固有空間へと導ける。
- (性質2) W の最小固有値をもつ固有ベクトル $e^{(min)}$ の固有値を V では 0 にしているため、線形変換でその成分は一度完全に消去される。その後非線形変換が施されるが、これはローカルにしか働かないため、状態ベクトルは $e^{(min)}$ の成分をほとんど持たない。他の負の固有ベクトルも正の大きな固有ベクトルに比べてその絶対値は小さく、その成分を大きくは持たなくなる。

このように、行列 V は上で条件としてあげた二点を満足しており、この行列を使えば離散時間同期型ダイナミクスでも準最適解の探索が行えると期待できる。実際、この行列を用いて離散時間同期型ダイナミクスによる解探索を行った時のエネルギーの減少過程 ($N = 500$) を図 8.5 に示す。但し、ここでは組み合わせ最適化問題の解としての評価を見るため、正規のリアプノフ関数 $E(x)$ ではなく $E(y)$ の方をプロットしている。また、比較の対象としては連続時間ダイナミクスを模した差分方程式モデルで解探索を最速に行える差分を用いた場合の計算結果を併記している。この図が示すように、行列 V を用

いた離散時間同期型ダイナミクスで解探索を行うと、計算の繰り返しが少なくてもエネルギーが減少し、良質の解が得られている。

ここで、次のような疑問が想定される。それは、上で提案した手法は行列の成分のうち制約条件違反に対するペナルティ項の固有値成分を取り除いているのであるから、制約条件のペナルティの係数を小さくして行列を作ればよいのではという疑問である。しかし、この方法はうまく機能しない。なぜなら、この方法では固有ベクトルが変わってしまうからである。つまり、制約条件を満たさない空間に小さい固有値を持つ固有ベクトルがあてがわれなくなってしまう。その結果、制約条件を満たさない解に収束しやすくなる。ここで提案した方法の要点は、固有ベクトルを変えずに絶対値の大きい負の固有値の抑制だけを行っている点である。これが目標とする準最適解に収束先を維持しつつ高速計算を行うことを可能にしているのである。

8.3 巡回セールスマン問題の幾何と応用

分割問題の場合は $(-1, 1)$ 出力型のニューロンで解の候補が表現されるので、前節の連想記憶の延長で議論を行うことができた。しかし、実際は $(0, 1)$ 出力型のニューロンで表現される組み合わせ最適化問題の方がより一般的である。そこで、本小節では分割問題と同じ手法を $(0, 1)$ 出力系で表される巡回セールスマン問題 (Traveling Salesman Problem, TSP) に応用することを検討する。

巡回セールスマン問題 (TSP) とは、 N 個の都市を一度ずつ全部廻る最短経路を探す問題である (図 8.6)。今、都市 a と都市 b の距離を d_{ab} とする。また、ニューロンを N^2 個用意し、 (a, i) ニューロンの出力 x_{ai} が 1 なら時刻 i に都市 a を通ることとする。すると、最短にすべき経路の長さは

$$C(\mathbf{x}) = \frac{1}{2} \sum_a \sum_{b \neq a} \sum_i d_{ab} x_{ai} (x_{b,i+1} + x_{b,i-1}) \quad (8.26)$$

となる。これがコストの項になる。一方、制約条件は、各都市を一度ずつしか通らないこと、同じ時刻に 2 つ以上の都市を廻れないこと、全部で N 都市訪問することの 3 つである。これらの制約を破ることのペナルティは

$$P(\mathbf{x}) = \frac{A}{2} \sum_a \sum_i \sum_{j \neq i} x_{ai} x_{aj} + \frac{B}{2} \sum_i \sum_a \sum_{b \neq a} x_{ai} x_{bi} + \frac{C}{2} (\sum_a \sum_i x_{ai} - N)^2 \quad (8.27)$$

と表現できる。このコスト項とペナルティ項の和

$$E(\mathbf{x}) = C(\mathbf{x}) + DP(\mathbf{x}) \quad (8.28)$$

が系のエネルギーとなるように荷重行列を設定すれば良い (A, B, C, D は定数)。但し、このままではコストを表す結合荷重が全て負になり、その結果平均発火率が低下し制約

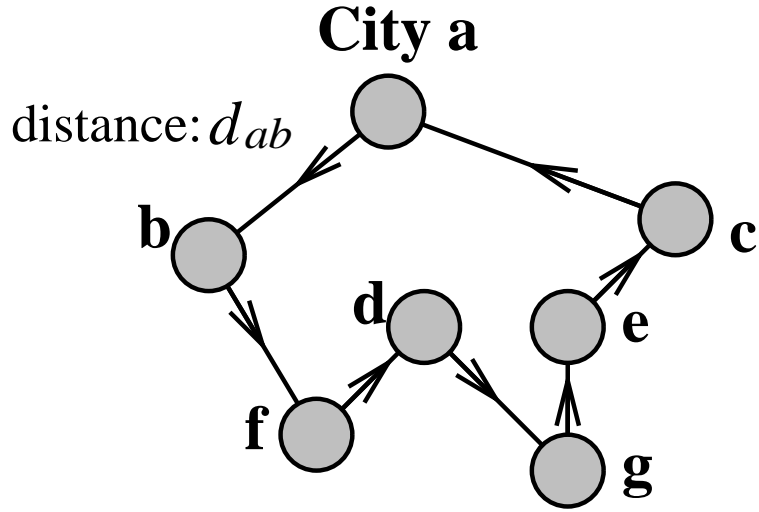


図 8.6: 巡回セールスマン問題。

条件を満たさない解を生じやすくなる。Hopfieldらはこの問題を制約条件の一つを表す項 $\frac{C}{2}(\sum_a \sum_i x_{ai} - N)^2$ を $\frac{C}{2}(\sum_a \sum_i x_{ai} - \kappa N)^2 (\kappa > 1)$ とすることで回避していたが、本論文ではコスト項の距離 d_{ab} を $d'_{ab} = d_{ab} - h (h > 0)$ に置き換えることで制約条件を満たす解を得やすくする。

(a, i) ニューロンと (b, j) ニューロンの結合を w_{ajib} 、 (a, i) ニューロンの閾値を θ_{ai} とすると、上記のエネルギー関数を満たす w_{ajib}, θ_{ai} はそれぞれ

$$w_{ajib} = -A\delta_{ab}(1 - \delta_{ij}) - B\delta_{ij}(1 - \delta_{ab}) - C - Dd'_{ab}(\delta_{j,i+1} + \delta_{j,i-1}) \tag{8.29}$$

$$\theta_{ai} = -CN \tag{8.30}$$

となる。そして、ダイナミクスとして

$$\dot{u}_{ai} = -\frac{1}{\tau}u_{ai} + \sum_b \sum_j w_{ajib}x_{bj} \tag{8.31}$$

$$x_{ai} = f(u_{ai} - \theta_{ai}) \tag{8.32}$$

$$f(u) = \frac{1}{(1 + \exp(-\beta u))} \tag{8.33}$$

を用いる。デジタル計算機を使う場合は、式 (8.31) を差分方程式

$$u_{ai}(t + \Delta) = (1 - \Delta)u_{ai}(t) + \Delta \sum_b \sum_j w_{ajib}x_{bj}(t) \tag{8.34}$$

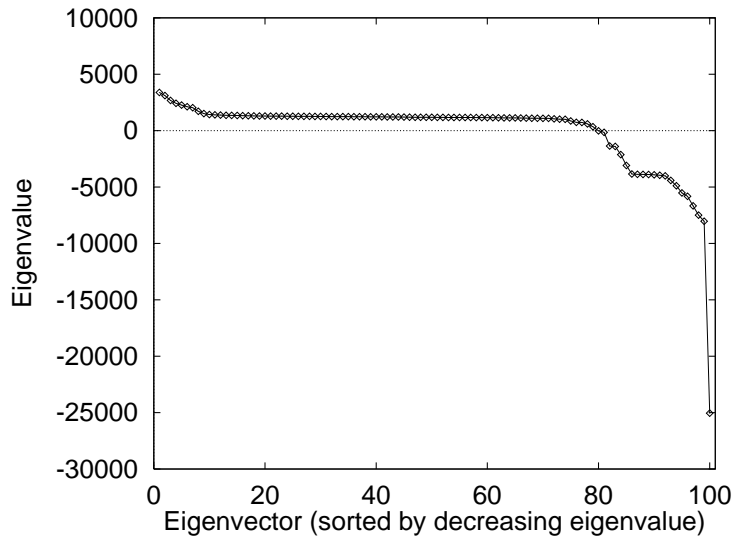


図 8.7: TSP を解くネットワークの荷重行列の固有値。

に置き換えて計算することになる。この計算ならば並列同期型のデジタル計算機で実行することが可能である。しかし、ここでも、差分 Δ が小さいと収束までに必要な計算が多くなり、逆に Δ が大きいと系が振動的になり解の探索に失敗する。

ここでは、Hopfieldらの論文[18]中にある TSP の例題についてその幾何構造を解析し、その解析結果に基づき解の並列高速探索法を提案する。以下の議論では、前章で x_{ai} と記したニューロンの番号を 1 次元に並べ換え x_i と表現する。同様に結合荷重や閾値も w_{ij}, θ_i と書く。文献[18]においては、 $[0, 1] \times [0, 1]$ の二次元空間にランダムに配置された 10 都市を巡回する最短経路を求める例題が取り上げられている。また、荷重行列構成時のパラメータは $A = 500, B = 500, C = 200, D = 500$ と設定されている。この時に構成される荷重行列の固有値の分布を図 8.7 に示す。但し、この論文で導入したパラメータ h は 0.8 に設定している。図 8.7 から分かるように、TSP でも分割問題の場合と同様、突出した最小固有値が存在している。

分割問題では、絶対値の大きい負の固有値成分を取り除くことで系の振動を抑え、大きい差分 Δ を使って解を高速探索することが可能であった。よって、TSP においても、分割問題と同様にして、並列同期計算における解の高速探索ができると期待される。しかしながら、 $(0, 1)$ ニューロンモデルの場合は

$$v_{ij} = w_{ij} - c\lambda_{\min}^{(i)} e_i^{(i)} e_j^{(j)} \quad (8.35)$$

のように最小固有値成分を $c (< 1)$ 倍に減じるだけでは、ニューロンの発火率が急上昇し、制約条件を満たさない解が現れてしまう。これを防ぐためには、閾値を適切な値まで上

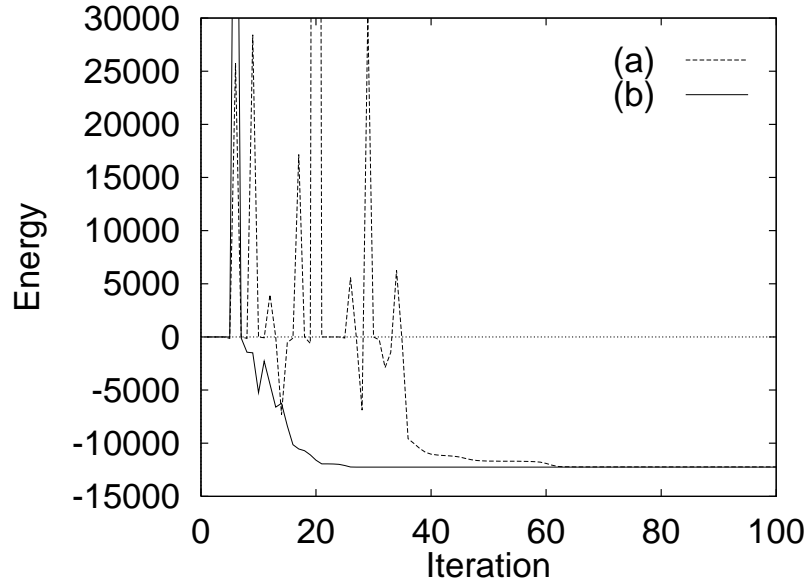


図 8.8: エネルギー減少過程の比較 ($\beta = 0.1$): (a) 従来の行列を用いた場合 ($c = 0, \Delta = 0.03$); (b) 行列と閾値を変更して時間差分を大きくとった場合 ($c = 0.7, \Delta = 0.28$)。

げる必要がある。閾値の上げ幅は、ニューロンの発火率が $1/N$ であることを考慮し、

$$\theta_i = -CN - \frac{c\lambda_{min}}{N} \sum_j e_i^{(min)} e_j^{(min)} \quad (8.36)$$

とすれば、発火率を変えずに最小固有値成分を取り除き、時間刻み幅 Δ が大きくても準最適解が得られるようになると予想される。図 8.8 に従来のモデルと最小固有値成分を取り除くと同時に閾値も上記の通り修正したモデルのエネルギー減少過程を比較した結果を示す。このように、 $(0, 1)$ ニューロンの場合でも、最小固有値成分を取り除くと同時に閾値を修正すれば、大きい差分を用いてもエネルギーが減少し、準最適解を高速で得ることができる。

以上のように、TSP のような $(0, 1)$ 出力型ニューロンで表現される組み合わせ最適化問題においても、固有空間の修正によって並列同期計算を利用した解探索の高速アルゴリズムを構成することができた。

8.4 まとめ

本章では、ニューロダイナミクスを幾何学的な視点から捉えることにより、組み合わせ最適化問題において、並列デジタル計算機を使ってと高速に準最適解を得るアルゴリズムを提案した。提案した手法では、制約条件に対応する固有空間の固有値成分を取り除くことで、離散時間同期型ダイナミクスでも振動を抑え、高速に準最適解を探索することを可能にしている。提案したアルゴリズムの有効性を示すため、 $(0, 1)$ ニューロンで表現される問題の例とし分割問題、 $(-1, 1)$ ニューロンで表現される問題の例として巡回セールスマン問題を取り上げ、どちらの問題においても高速に準最適解が得られることを確かめた。

今後の課題としては、ローカルミニマムへの停留を防ぐ方法として提案されている種々のアルゴリズムと、ここで提案したアルゴリズムを組み合わせ、より良質な解への高速収束を実現する方法を検討することが挙げられる。

第 9 章： 学習に関する考察

前章までの議論においては、学習に関しては Hebb 型の学習のみを考えていた。しかし、ニューラルネットワークにおける学習には他にも多くの学習法が知られている。本章においては、それらの学習法にまつわる話題について議論を行う。

バックプロパゲーション (BP) 等の繰り返し学習を考えると、ニューラルネットワークのダイナミクスとして、二種類のダイナミクスを想定することができる。一つは、前章まででも議論したニューロンの状態遷移のダイナミクスであり、もう一つは荷重行列が変化するダイナミクスである。であるから、幾何的な解析を論じる場合、学習の結果できあがったネットワークの固有空間解析による系の力学的性質の他に、学習過程における荷重行列の固有空間の構造の変化という議論も行うことができる。よって、幾何的な解析の応用範囲の可能性もそれだけ広がることになる。

本章は、学習に関する初歩的な考察として、直交学習と BP 学習を取り上げる。直交学習については、学習の結果得られたネットワークの幾何構造を論じ、一方 BP 学習については学習過程のダイナミクスについて議論を行うことにする。

9.1 直交学習に関する考察

9.1.1 連想記憶における直交学習

本論文のこれまでの連想記憶に関する議論では、全て Hebb 則に基づく相関学習、或はコバリアンس学習を用いていた。連想記憶を実現するモデルとして、他に直交学習

$$w_{ij}^s(t+1) = (1 - \tau)w_{ij}^s(t) + c \left(\sum_{k=1}^N w_{ik}^s s_k - s_i \right) s_j \quad (9.1)$$

で与えられる [3]。ここで、 τ と c は十分小さい定数である。

この直交学習を繰り返すと、記憶行列は

$$W^s = \sum_{\mu=1}^P \mathbf{s}^{\mu} \mathbf{s}^{\mu*} \quad (9.2)$$

に収束する。ここで、 $\mathbf{s}^{\mu*}$ は記憶パターンベクトル \mathbf{s}^{μ} の双対ベクトル

$$\mathbf{s}^{\mu*} \mathbf{s}^{\nu} = \delta_{\mu\nu} \quad (9.3)$$

である。この連想記憶モデルを一般逆行列モデルと呼ぶ。

この行列は、自己相関記憶行列と違って、一般には対称行列ではない。記憶行列 W^s の固有ベクトルは、 $P < N$ ならば、記憶パターン \mathbf{s}^{μ} に完全に一致し、その固有値は 1 で

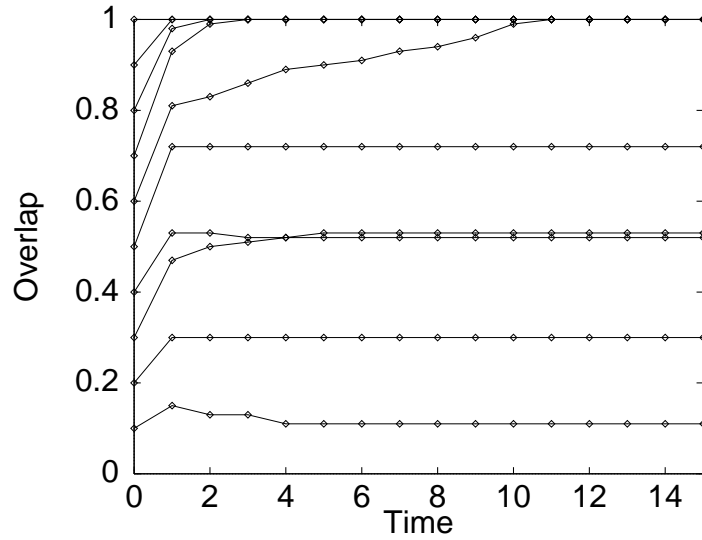
ある。一方、ノイズ空間の固有値は0である。この固有空間構造から、D関数は記憶パターンで1となり、それ以外では1以下になることが分かる。つまり、自己相関連想記憶のように、記憶パターンよりもD関数が大きな値となる偽記憶は生じない。よって、記憶したパターン数が N までなら、記憶パターンは安定になることがわかる。また、ノイズ空間の固有値は0になっていることから、自己相関連想記憶の時と同様に、ノイズ成分は線形変換で打ち消され、パターンは記憶したパターンへと引き込まれていくことになる。

しかしながら、実際は、 N 個までパターンを安定に記憶できても、 $P = 0.5N$ の時点で引き込み領域が失われ、連想記憶としての機能が果たされなくなる。これは、一般逆行列モデルの記憶行列の対角成分 w_{ii}^s が平均して a となるので、もし a が 0.5 を越えると記憶パターンに近いパターン $\mathbf{x} \sim \mathbf{s}$ について

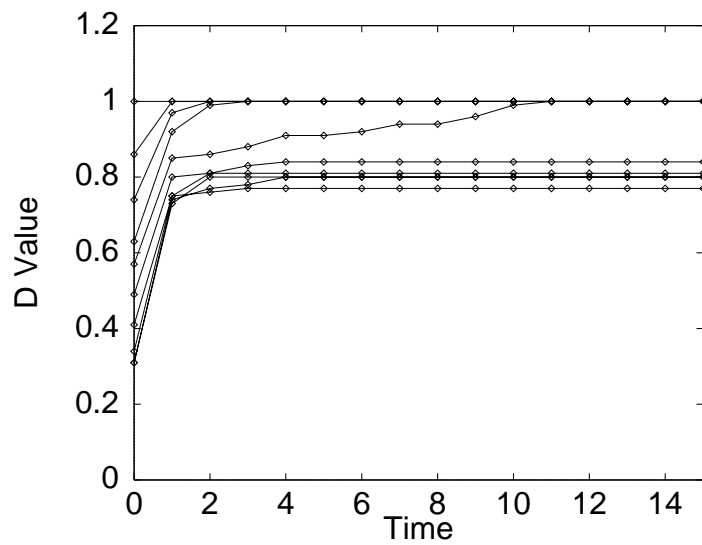
$$\left| \sum_{j \neq i} w_{ij}^s x_j \right| < |w_{ii}^s| \quad (9.4)$$

が成り立つことによる。つまり、記憶パターンの付近で、自己結合によって引き込みが妨げられてしまうことになる。 W^s を用いた連想記憶の想起過程におけるD関数の振舞いを図9.1(記憶率0.3)、図9.2(記憶率0.6)に挙げる。このように、記憶率を0.5以上にすると、 W^s を用いたモデルでは、D関数が1になるまでに止まってしまう。

上で述べたことから、記憶行列 W^s から自己結合を取り除いた行列 W^0 を記憶行列として用いると、引き込み領域が多少増大すると予想される。このモデルについては次小節の想起法の改良に関する議論で触れることにする。

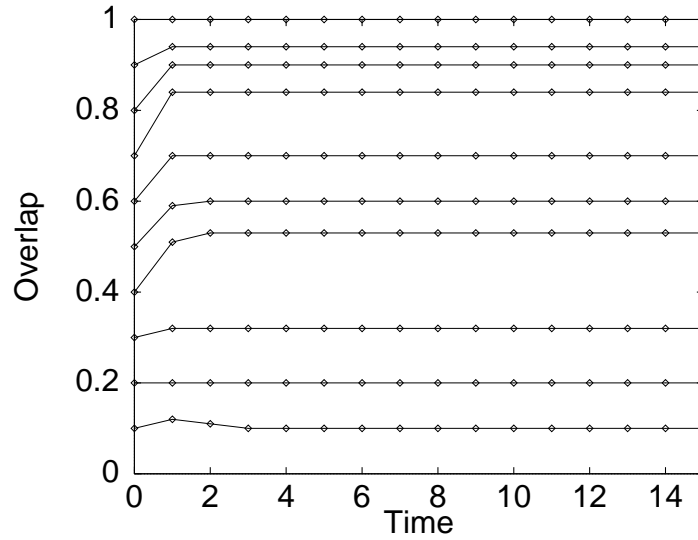


(a) オーバーラップの変化



(b) D 関数の変化

図 9.1: 一般逆行列モデルにおけるオーバーラップと D 関数のダイナミクス (自己結合あり, $a = 0.3, N = 200$)。



(a) オーバーラップの変化

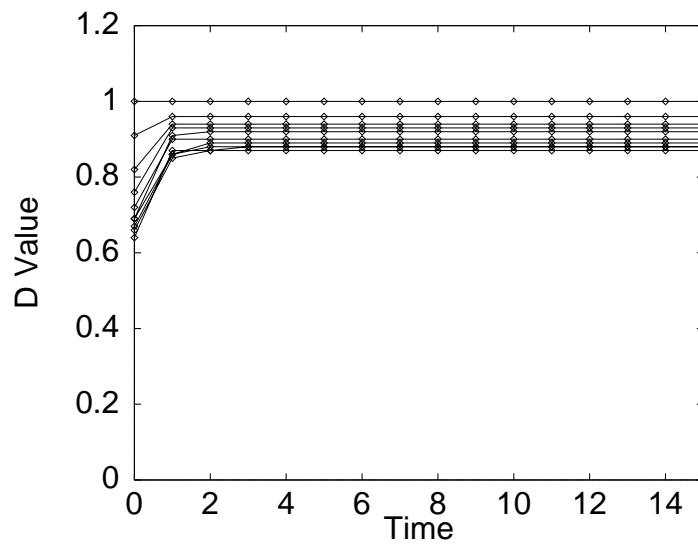
(b) D 関数の変化

図 9.2: 一般逆行列モデルにおけるオーバーラップと D 関数のダイナミクス (自己結合あり, $a = 0.6$, $N = 200$)。

9.1.2 一般逆行列モデルの想起法改良

部分反転法は一般逆行列モデルにおける想起法としても、引き込み領域を拡大する効果が確認されている [28]。ここでは、連想記憶の幾何学的解析手法の視点から、一般逆行列モデルの引き込み領域を更に拡大する想起ダイナミクスを提案する。

上で述べたように、一般逆行列モデルを用いた時は、D 関数が 1 になる前に、ダイナミクスが停止してしまう。であるから、この停止状態を抜け出させるような想起ダイナミクスを考えれば良い。

今、ダイナミクスの改良を考えるために、

$$\mathbf{u}^\mu = W \mathbf{s}^\mu \quad (9.5)$$

という量を定義する。これは、記憶パターンを入力し、それに荷重行列をかけた時に得られるベクトルである。

一般逆行列 W^s を記憶行列として用いた時、記憶パターン \mathbf{s}^μ は固有値 1 の固有ベクトルになる。つまり、 $|u_i^\mu| = 1$ が全ての μ と i について成り立つわけである。このことから、記憶パターンだけを安定化するには、内部状態が 1 付近のニューロンだけを安定化し、他のニューロンを不安定化すればよい。これは、部分反転法のパラメータを変えることによって簡単に行うことができる。

いくつかの想起ダイナミクスを用いて、一般逆行列モデルの記憶パターンを想起した数値実験の結果を図 9.4 に示す。まず、内部状態 1 付近のニューロンだけを安定化し、他を不安定化する方法として、図 9.3(A) に示した関数を $\phi(u)$ に用いると、記憶率 a が大きい時は引き込み領域が大きくなるが、 a が小さい時は引き込み領域が小さくなる。これは、ノイズ空間の次元が高い時は、小さな内部状態をもつニューロンを不安定化することで、記憶空間への引き込みを阻害しているからであると考えられる。図 9.3(B) に示した $\phi(u)$ を使うと、記憶率 a が小さい時も、引き込み領域が大きく保たれるが、記憶率 a が大きい時には、先程のモデルよりも引き込み領域が小さくなる。記憶率 a の大小に関わらず相対的に広い引き込み領域を確保するには、次のような 2 つのフェーズをもつ想起方法が有効である。まず、第 1 のフェーズでは、普通の符号関数を非線形関数として想起することによって、ノイズ空間の成分を落とす。次に、第 2 フェーズとして図 9.3(B) に示した $\phi(u)$ を使った部分反転法を用いて想起することで、記憶パターンへの収束を促す。この方法により、記憶率 a が小さい時も大きい時も、優秀な連想性が保たれる。この方法を使うと、記憶率 0.8 付近まで引き込み領域が記憶率の増加に比例して線形に減少するので、特に記憶率が 0.4 から 0.8 付近で自己結合をとったモデルよりも引き込み領域が広がる。また、全小節で述べた通り、ダイナミクスを変更しなくても、自己結合を除くことで引き込み領域の増加が図れる。しかし、このモデルも上記で示した中で最適な改良ダイナミクスモデルと比較すると引き込み領域が小さくなっている。

最後にこの 2 つのフェーズをもったダイナミクスで想起した時のオーバーラップと D 関数の時間変化を図 9.5、図 9.6 に示す。これらの図から、第 2 フェーズで図 9.3(B) に示した絶対値の小さい内部状態を不安定化するダイナミクスによって、パターンが記憶バンドの方へと押し上げられる様子がわかる。

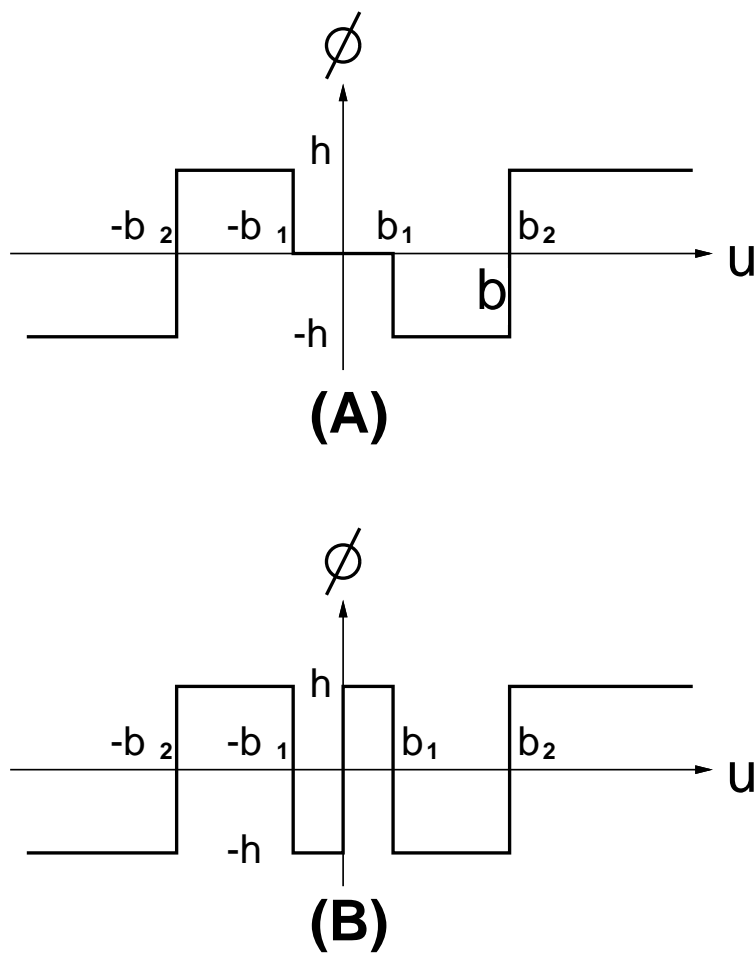


図 9.3: 一般逆行列モデルに応用した部分反転法: (A) $|u| = 1$ 付近を安定化し、 $|u|$ が大きいところを不安定化する関数 $\phi(u)$; (B) $|u| = 1$ 付近を安定化し、 $|u|$ が大きいところと小さいところを共に不安定化する関数 $\phi(u)$ 。

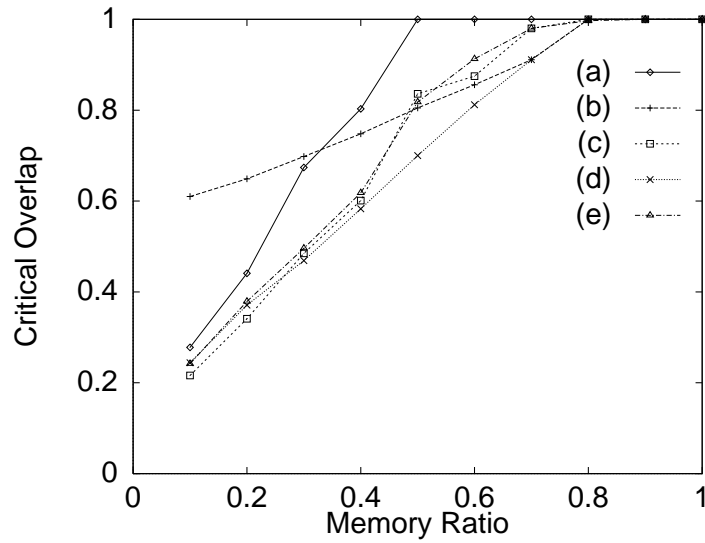
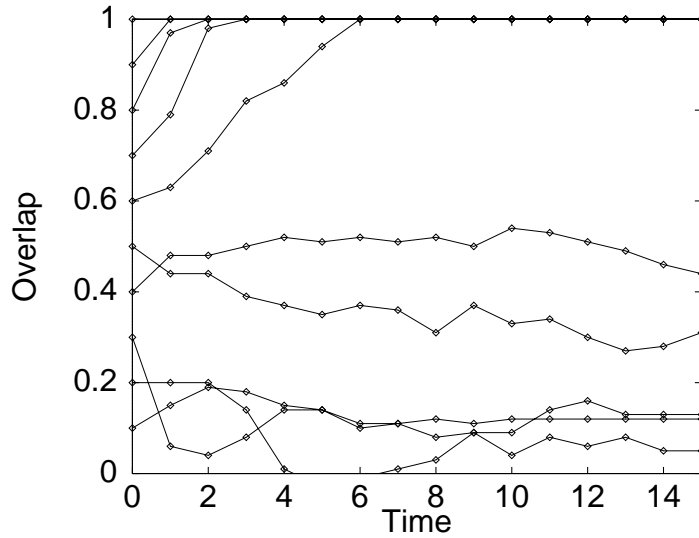
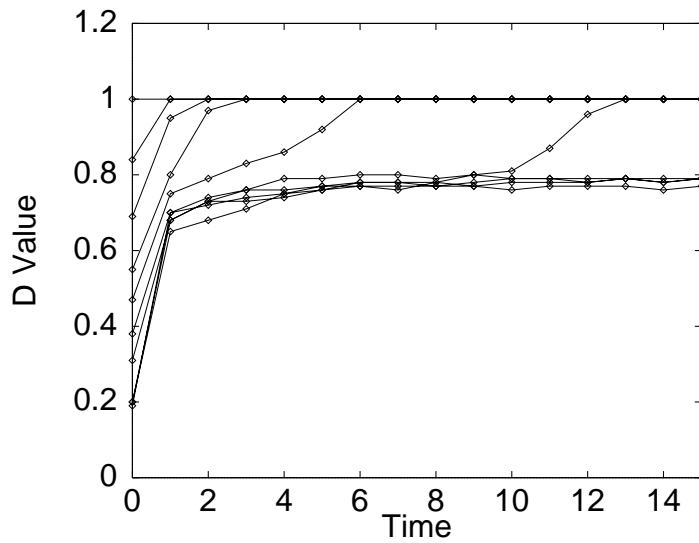


図 9.4: 一般逆行列モデルの引き込み領域 ($N=100$): (a) 通常の想起法; (b) 部分反転法 (A)($b_1 = 0.5, b_2 = 1.5, h = 1.2$); (c) 部分反転法 (B)($b_1 = 0.5, b_2 = 1.5, h = 1.2$); (d) 通常の想起法 (第1フェーズ)+部分反転法 (B)(第2フェーズ); (e) 自己結合を除いたモデル。

以上のように、直交学習で得られた一般逆行列による連想記憶モデルについて、その固有空間の幾何構造をもとに、その連想性の向上を計る想起ダイナミクスを得ることができた。

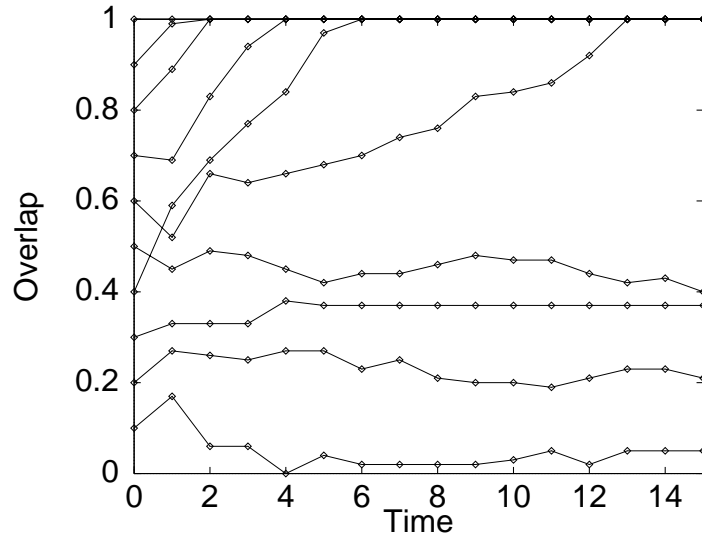


(a) オーバーラップの変化

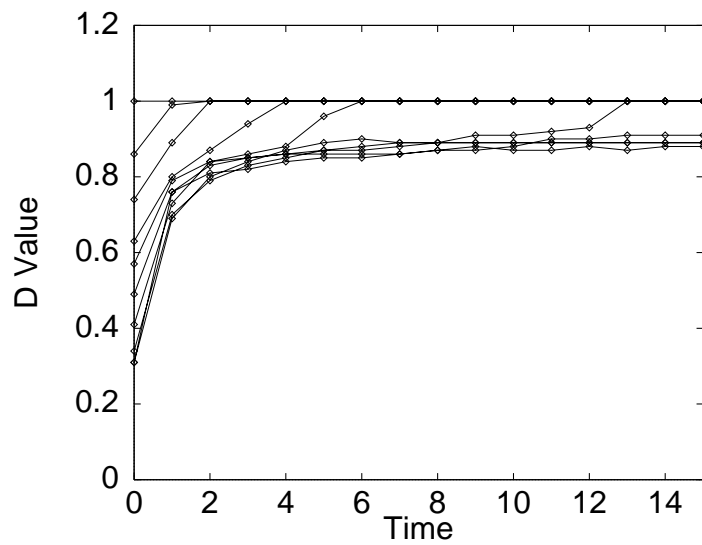


(b) D 関数の変化

図 9.5: 一般逆行列モデルにおいて最初のステップでは通常の想起、その後は部分反転法で想起した時のオーバーラップと D 関数のダイナミクス ($a = 0.3, N = 200$)。



(a) オーバーラップの変化



(b) D 関数の変化

図 9.6: 一般逆行列モデルにおいて最初のステップでは通常の想起、その後は部分反転法で想起した時のオーバーラップと D 関数のダイナミクス ($a = 0.6, N = 200$)。

9.2 BP学習に関する考察

多層パーセプトロンにおけるBP学習では、適切な中間層の情報表現を自動生成することにより、多次元の複雑な関数を実現している。しかし、中間層の数や学習定数を事前に適切な値に設定しなければ、BP学習は成功しない。

BP学習については、今までにも様々な研究が行われてきた。特に、関数近似の方法としてのBP学習については非常に多くの研究がなされている。しかし、現在に至るまで、BP学習を完全に信頼のおける学習として使用できるだけの理論的基盤は十分に整備されていない。この節では、論理関数近似におけるBP学習について、中間層の情報表現に着目しながら、その学習のダイナミクスの特性を考察する。

9.2.1 BP学習

パーセプトロンにおいては、入力 \mathbf{x} と出力 \mathbf{y} の入出力関係は

$$y_i = f(u_i) \quad (9.6)$$

$$= f\left(\sum_j w_{ij}x_j + \theta\right) \quad (9.7)$$

で与えられる。ここで、関数 $f(u)$ は非線形関数で、その出力を 0 以上 1 以下とする時は、

$$f_0(u) = \frac{1}{1 + \exp(-u/T)} \quad (9.8)$$

-1 以上 1 以下とする時は

$$f_1(u) = \frac{1 - \exp(-u/T)}{1 + \exp(-u/T)} \quad (9.9)$$

がよく使われる。以下でも、これらをニューロンの活性化関数として用いる。これを L 層の入出力 \mathbf{x}^μ をもつ、

$$x_i^{(\mu+1)} = f(u_i^\mu) \quad (9.10)$$

$$= f\left(\sum_j w_{ij}^{(\mu, \mu+1)}x_j^{(\mu)} + \theta_i^{(\mu+1)}\right) \quad (9.11)$$

($\mu = 1, \dots, L-1$) で与えられるシステムに拡張することができる。以下では、3層のパーセプトロンを考える。また、簡単のため、全ての層は N 個のニューロンから成るとする。

BP学習とは、最急降下法を多層パーセプトロンに応用したものである。教師信号を ζ 、学習定数を c とすると、3層パーセプトロンのBP学習は、

$$dw_{ij}^{(2,3)} = c(\zeta_i - x_i^{(3)})f'(u_i^{(3)})x_j^{(2)} \quad (9.12)$$

$$d\theta_i^{(3)} = c(\zeta_i - x_i^{(3)})f'(u_i^{(3)}) \quad (9.13)$$

$$dw_{ij}^{(1,2)} = c \sum_k (\zeta_k - x_k^{(3)}) f'(u_k^{(3)}) w_{ki}^{(2,3)} f'(u_i^{(2)}) x_j^{(1)} \quad (9.14)$$

$$d\theta_i^{(2)} = c \sum_k (\zeta_k - x_k^{(3)}) f'(u_k^{(3)}) w_{ki}^{(2,3)} f'(u_i^{(2)}) \quad (9.15)$$

で与えられる。ここで、関数 $f(u)$ として $f_0(u)$ を用いた場合は

$$f'_0(u) = f_0(u)(1 - f_0(u)) \quad (9.16)$$

$f_1(u)$ を用いた場合は

$$f'_1(u) = 0.5(1 + f_1(u))(1 - f_1(u)) \quad (9.17)$$

である。

9.2.2 BP 学習による中間層発火率の変化

ここでは、 N 次元ベクトル $\boldsymbol{\xi}^p = (\xi_1^p, \xi_2^p, \dots, \xi_N^p)$, $\boldsymbol{\zeta}^p = (\zeta_1^p, \zeta_2^p, \dots, \zeta_N^p)$ をそれぞれ P 個ずつ用意し、そのペアが入出力関係 $\zeta_i^p = F_i(\boldsymbol{\xi}^p)$ を満たす論理関数 $\mathbf{F} = (F_1, F_2, \dots, F_N)$ を、入力層、中間層、出力層が全て N 個のニューロンからなる 3 層パーセプトロンにバックプロパゲーションを使って学習させる。入出力ベクトルについては、成分が -1 と 1 の二値をとる場合と 0 と 1 の二値をとる場合を考える。 -1 と 1 の二値論理の場合はニューロンの活性化関数として $f_1(u)$ を採用し、 0 と 1 の二値論理の場合はニューロンの活性化関数として $f_0(u)$ を採用することにする。

数値実験として、 $N = 50$ のパーセプトロンに、 $P = 100$ 個の入出力関係を学習定数 $c = 0.1$ で学習させた場合の中間層ニューロンの発火率の変化と出力誤差の変化を調べる。図 9.7 に $-1/1$ モデル、図 9.8 に $0/1$ モデルにおける結果を示す。ここでは、学習は $T = 1$ で行い、評価は中間層の発火率、誤差共に $T = \infty$ として行う。両モデルを平等に扱うために、誤差は

$$E = \frac{1}{P} \sum_{p=1}^P \sum_{i=1}^N (1 - \delta(\zeta_i^p - F_i(\boldsymbol{\xi}^p))) \quad (9.18)$$

$$\delta(u) = \begin{cases} 1, & \text{if } u = 0 \\ 0, & \text{if } u < 0 \end{cases} \quad (9.19)$$

で定義する。

図 9.7、図 9.8 の実験結果からわかるように、 $-1/1$ モデルでは中間層の発火率は大きく変化しないが、 $0/1$ モデルでは学習の初期段階で中間層のニューロンの発火頻度が低下していく現象が見られる。また、図 9.9、図 9.10 の様に、その低下は学習するデータの数 P と学習定数 c が増加するほど大きくなる。

9.2.3 考察

0/1モデルにおける中間層発火率低下のメカニズムは、定性的には以下のように説明できる(図9.11)。

中間層の発火頻度を決定するのは第1層から第2層への結合 $w_{ij}^{(1,2)}$ と、第2層のニューロンの閾値 $\theta_i^{(2)}$ であるから、これらのパラメータの学習則に注目してみることにする。以下では、簡単のため、中間層、出力層が0と1だけを出力すると考えて議論を進める。

まず、 $w_{ij}^{(1,2)}$ に着目する。このパラメータの学習は式(9.14)によって与えられる。この式から分かるように、学習が進むのは入力 x_j が1で、出力が教師と異なる場合に限られる。また、 $f'(u_k^{(3)})$ と $f'(u_i^{(2)})$ は常に正であるから、学習が正に進むか負に進むかは $(\zeta_k - x_k^{(3)})w_{ki}^{(2,3)}$ の符号によって決まる。つまり、中間層から出力層への結合 $w_{ki}^{(2,3)}$ の正負、出力 $x_k^{(3)}$ の0/1によって4通りの場合に分けると、 $w_{ij}^{(1,2)}$ が増加するように学習が進むのが、 $w_{ki}^{(2,3)} > 0, x_k^{(3)} = 0$ の場合と $w_{ki}^{(2,3)} < 0, x_k^{(3)} = 1$ の場合、 $w_{ij}^{(1,2)}$ が減少するように学習が進むのが $w_{ki}^{(2,3)} < 0, x_k^{(3)} = 0$ の場合と $w_{ki}^{(2,3)} > 0, x_k^{(3)} = 1$ の場合である。ここで、中間層が1を出力している場合は、 $w_{ki}^{(2,3)}$ と $x_k^{(3)}$ には相関がある。 $w_{ki}^{(2,3)} > 0$ なら、 $x_k^{(3)}$ は1に、 $w_{ki}^{(2,3)} < 0$ なら、 $x_k^{(3)}$ は0になりやすくなる。つまり、一般に $w_{ij}^{(1,2)}$ が減少するように学習が進む場合の方が多くなる。もちろん、この説明が成り立つのは、入出力や結合荷重がランダム性を持つ学習の初期段階に限られる。

$\theta_i^{(2)}$ は、常に入力が1の $w_{ij}^{(1,2)}$ と見なすことができる。この場合は、 $\theta_i^{(2)}$ が減少する方向に常にバイアスがかかっていることになり、その減少のスピードは2倍になっている。しかし、 N 個のニューロンからの結合 $w_{ij}^{(1,2)}$ は全て負のバイアスを持つことを考えると、それでも閾値が高すぎて、中間層の発火率は低く抑えられることになる。ここで、入力の発火率が50%で閾値の学習が2倍であることを考慮すると、閾値が $N/4$ 倍のゲインを持てば、中間層の発火率の減少は抑制できる計算になる。実際、このように閾値を設定して数値実験を行うと、図9.12に示す通り発火率の低下は抑制されることが確認できる。また、逆に、閾値のゲインを小さく保つと、図9.13の様に発火率の低下は促進される。

以上で行った議論は、BP学習のダイナミクスの初歩的解析である。この時点では、まだ固有空間解析等の幾何的な解析手法は全く行っていない。今後は、学習過程における固有空間の構造の変化等の解析を行うことで、ニューロンの学習のメカニズムを調べるのが課題となる。

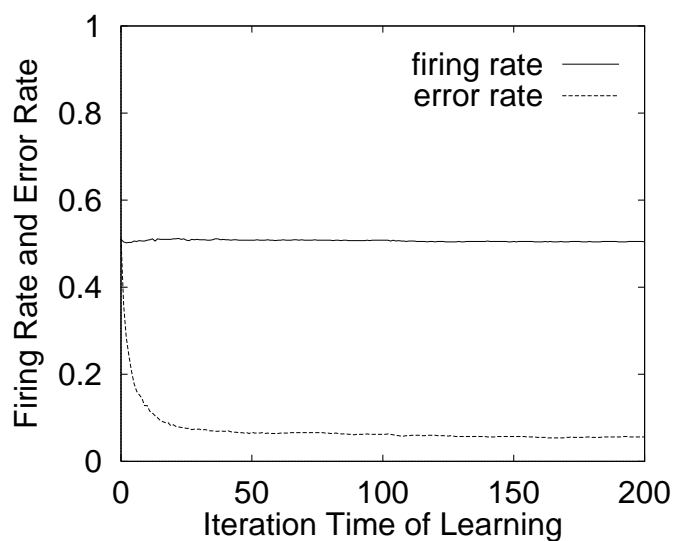


図 9.7: $-1/1$ モデルにおける中間層発火率と誤差の変化 ($c = 0.1, P = 100$)。

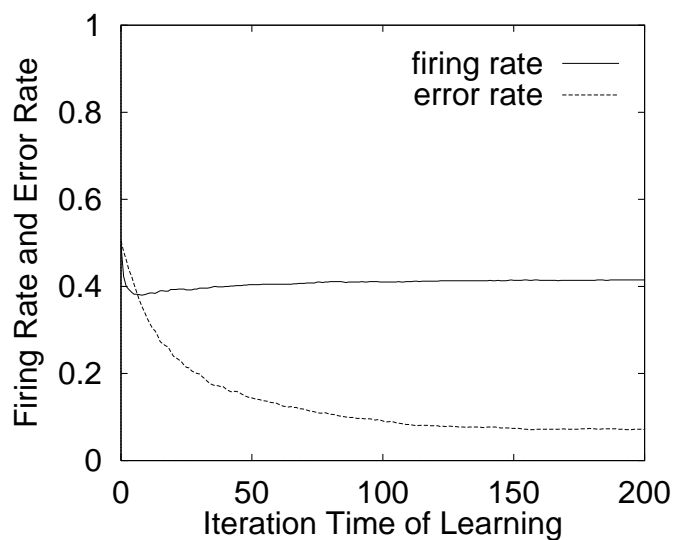


図 9.8: $0/1$ モデルにおける中間層発火率と誤差の変化 ($c = 0.1, P = 100$)。

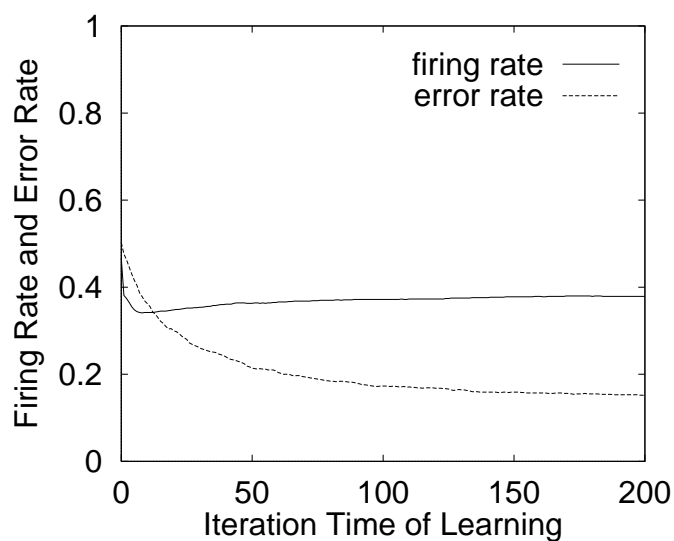


図 9.9: 0/1 モデルにおける中間層発火率と誤差の変化 ($c = 0.1, P = 200$)。

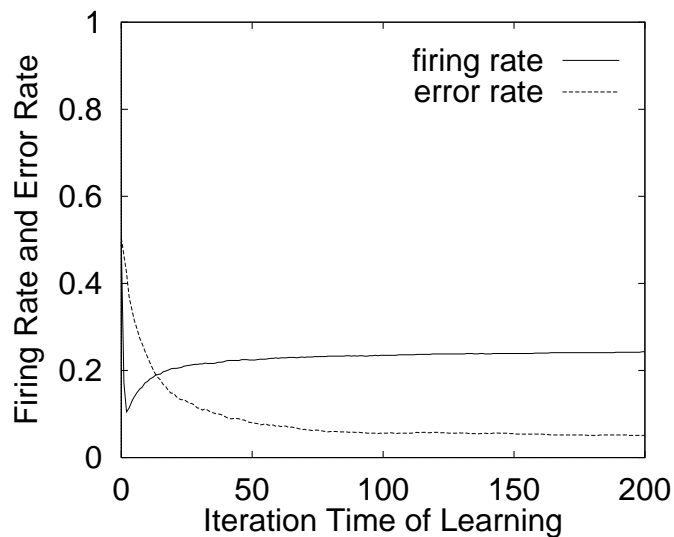


図 9.10: 0/1 モデルにおける中間層発火率と誤差の変化 ($c = 0.4, P = 100$)。

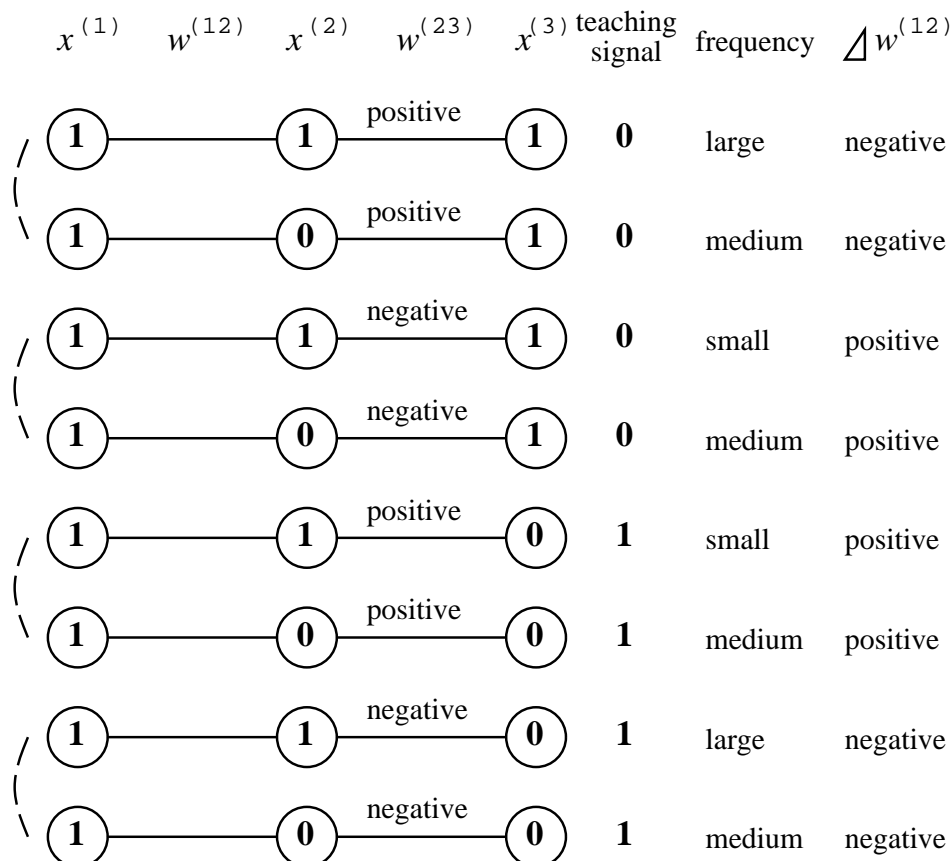


図 9.11: 中間層発火率低下のメカニズム。中間層の値は入力層と中間層の間の結合荷重 $w_{ij}^{(1,2)}$ に依存する。この $w_{ij}^{(1,2)}$ の学習が進むのは教師と出力の値が異なり、尚且つ入力 $x^{(1)}$ が 1 の時である。また、 $f'(u_k^{(3)})$ と $f'(u_i^{(2)})$ は常に正であるから、学習が正に進むか負に進むかは $(\zeta_k - x_k^{(3)})w_{ki}^{(2,3)}$ の符号によって決まる。この図のように場合わけをして考えると、学習は負に進むことが多いことが分かる。ニューロン数が多ければ、結合の学習は閾値の学習より大きな影響を持つから、中間層ニューロンの入力 u はより負になり易く、中間層の発火率は減少することになる。

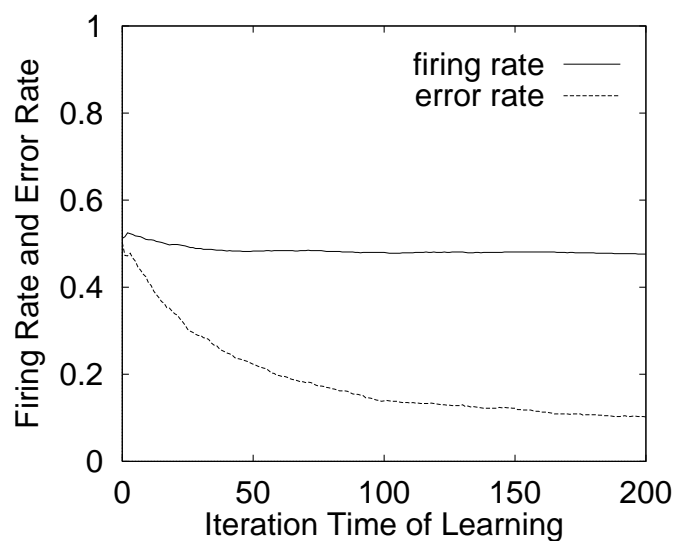


図 9.12: 閾値のゲインを $N/4 = 12.5$ 倍にしたときの中間層発火率と誤差の変化 ($c = 0.1, P = 100$)。

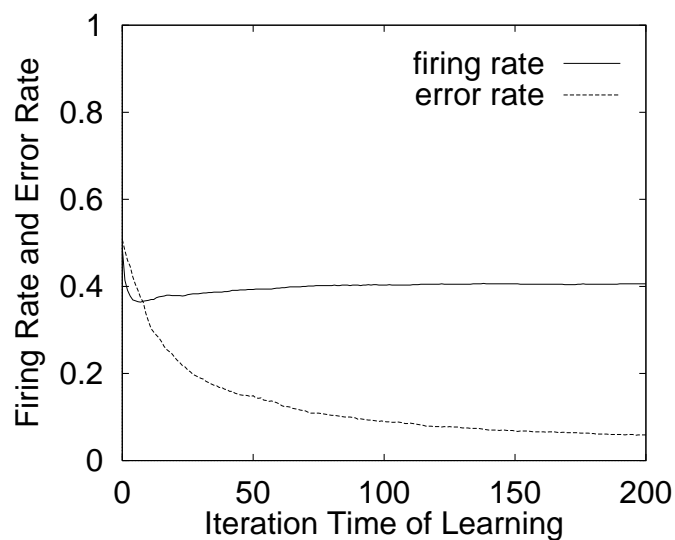


図 9.13: 閾値のゲインを 0.1 倍にしたときの中間層発火率と誤差の変化 ($c = 0.1, P = 100$)。

9.3 まとめ

本章では、直交学習、バックプロパゲーションなどの繰り返し学習を用いる場合について、得られたネットワークのダイナミクス特性、及び学習の過程における変化の様子について調べた。本章で行った研究は初歩的な議論で、幾何的解析法を用いた本格的議論を行うことはまだできていない。今後は、学習に関する幾何的議論を進めることで、リカレントネットワークによる制御系の学習過程、ロバスト性などの力学的性質を明らかにするとともに、新たな非線形力学制御法を開発していくことが課題となる。

第 10 章： 結論

10.1 本論文のまとめ

本論文では、ニューラルネットワークのダイナミクスにおいて線形変換がドミナントに働くことに着目し、荷重行列の固有空間解析に基づく幾何的議論で種々のニューラルネットワークモデルのダイナミクスが持つ性質を明らかにした。更に、その性質を応用してニューロダイナミクスを利用した新たな情報処理アルゴリズムの提案を行った。

本論文の議論を通して得られたニューロダイナミクスの描像は以下のように集約することができる。荷重付き配線で相互に結合した閾値素子のネットワークという形式を持つニューラルネットワークのダイナミクスにおいては、非線形効果が局所的であり、エネルギー関数の地形全体の概形を線形で近似できる (図 10.1)。一般の非線形問題では局所的な地形以上の情報が得にくいことと対比すると、ニューロダイナミクスは比較的取り組みやすい非線形システムであると言える。また、線形システムと比べると、それが最大固有値のみに支配されるのに比べ、ニューラルネットワークは固有値の分布が支配する力学系という点で、より多様な振る舞いを実現できる系であると言える。

この知見をもとに、ニューロダイナミクスを利用し、以下に挙げる新たな情報処理アルゴリズムを提案した。

まず、自己相関連想記憶において、ニューロウインドウ法による階層的な概念形成モデルを提案した。ニューロウインドウ法は、ニューラルネットワークのダイナミクスにおいては大きい固有値を持つ固有ベクトルが系を支配するという制限を取り払い、ある特定の範囲の大きさの固有値を持つ固有ベクトル方向を安定化することを可能にする。これを利用し、階層的なクラスターをなすパターン記憶において、各階層レベルにおける

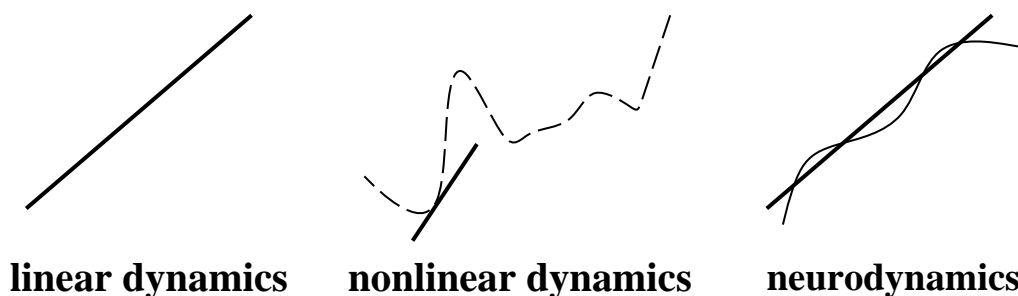


図 10.1: ニューラルネットワークのダイナミクスの特徴。全体の地形の概形を線形で近似できることがその扱いを容易にする。

代表パターンの選択的想起が実現された。

次に、非対称荷重を持つネットワークの幾何構造を解析することにより、相互関連想記憶において、ニューロウインドウ法を拡張してリミットサイクルの選択的取り出し法を提案した。

更に、組合せ最適化問題において、デジタル並列計算機で高速に準最適解を探索するアルゴリズムを提唱した。提唱した方法では、組合せ最適化問題を解くニューラルネットワークの結合荷重行列の固有空間の構造に着目し、計算を離散化した時に系の不安定化要因となる絶対値の大きな負の固有値成分を取り除いている。その結果、高速な収束を促す離散化が可能となり、準最適解の獲得までに必要な計算時間を格段に減らすことが可能となった。

このように、幾何的な解析によって得られたニューラルネットワークのダイナミクスの特徴、及びそれを応用した情報処理アルゴリズムをより定量的に解析するために、SCSNA、統計神経力学等の統計的な解析も同時に行った。それにより、提案したニューロウインドウ法の定量的な理論計算を実現し、パラメータ制御に対する理論的指針を与えることが可能となった。

10.2 今後の課題

本論文を終えるにあたり、最後に今後の課題を、本研究の延長線上にある課題、及びニューラルネットワーク研究の一般論としての課題に分けて以下にまとめる。

まず、本研究の延長線上にある課題を述べる。本研究においては、荷重行列の固有空間解析を使ってニューロダイナミクスの一般論の構築を試み、それを相関型連想記憶、組合せ最適化について応用した。しかし、ダイナミクスを持つニューラルネットワークとしては制御等に用いられるリカレントネットワーク等、他にも様々なものが存在している。それらについても、この理論を応用していくことで、非線形制御系の設計に有効な知見が得られる可能性はある。

また、ニューラルネットワークで使われる学習方法としても、本論文では主にHebb学習を論じたが、他にも種々の学習法が存在している。特に、繰り返し学習では、学習自体にもダイナミクスが存在することになる。第9章で繰り返し学習に関する予備的な議論は行ったが、応用可能なアルゴリズムの検討まで研究を進めるためには、幾何的解析手法等を導入することでその議論を深めていくこと必要となる。例えば、学習過程においてネットワークの荷重行列の変化を追うことができれば、新たな学習法の提案などへの指針が得られる可能性もある。

以上のような、幾何的な理論を応用する範囲の拡張の他にも、統計理論幾何理論と統計理論の関係に関する数理的研究、及び本論文で提案したニューロウインドウ法による階層概念形成回路を実際の生体の神経回路網でその存在の有無を確認することなどが本研究の延長線上で今後望まれる研究として挙げられる。

次に、ニューラルネットワーク研究の一般論としての課題を述べる。本研究では、既存の閾値型素子を荷重付き配線で組み合わせるという従来型のニューラルネットワーク

を前提に議論を進めた。その従来型のニューロで最も難しい部分の一つであるダイナミクスについて、その鳥瞰図を与えたことが本研究の意義であると言える。本研究によって、従来型の閾値型素子を荷重付きで配線したネットワークにおいては荷重行列の固有値の分布を利用した計算ができることが明らかにされた。しかし、逆に言うと、この事実はそれ以上のことはできないことも同時に意味している。であるから、実際の人間の脳が行っている種々の複雑な計算を可能にするには、生体脳が持つ、閾値素子や結合荷重といった特徴以上の新たな特徴抽出と定式化をしていく必要がある。今後のニューラルネットワーク研究の発展にはそのような研究が必要不可欠であろう。著者自身は、残念ながら、本論文でその仕事に手を付けることができなかったが、後進の優秀な研究者たちがその偉業を成し遂げることを祈りつつ、本論文を締めくくる次第である。

謝 辞

この研究の場を提供して頂いた指導教官の岡部洋一教授に感謝致します。岡部先生には、自由で恵まれた研究環境を与えて頂くとともに、研究内容に関しても熱心に議論の相手をして頂きました。特に、本論文の非対称ネットワークに関する研究においては、非対称行列の取り扱いについて貴重な助言を賜りました。また、研究テーマ以外でも、エントロピー、変分法といった数理的テーマから社会問題に至るまで、種々の議論の機会には何かと刺激を受けるところの多い三年間でした。

私が計数工学専攻在籍中にご指導頂いた現大阪大学基礎工学部教授の中野馨先生に感謝致します。中野研究室時代に培った研究者としての基盤は、私にとって貴重な財産となっております。修士課程の二年間、未熟な私を叱咤激励頂いた旧中野研究室の諸先輩方、及び東京大学工学部計数工学科の諸先生方にも、この場を借りて謝意を表したいと思います。

松下技研の金道敏樹先生には、私が修士課程に在学中、連想記憶の幾何学的解析手法という斬新で非常に興味深いテーマを紹介して頂き、本研究の基盤構築に多くのご指導、助言を賜りました。また、金道先生には学会発表や論文の書き方等、研究活動の基本も同時にご教授頂きました。ここにお礼申し上げます。

ERATOの岡田真人先生に感謝致します。岡田先生には、約二年間にわたり、本論文の統計的議論を進める上で多くの有益な助言を賜りました。

本論文を審査頂きました吉澤修治教授、合原一幸助教授、中須賀真一助教授、廣瀬明助教授に感謝致します。中でも、廣瀬先生は研究室の打ち合わせで、吉澤先生は学会や研究会等で、議論の相手をして頂く機会もよくありました。

北川助手、柴田元助手(現東工大)、宮尾技官、秘書の菅沼さん、白石さんには岡部研究室のスタッフとして、常日頃から研究のサポートをして頂きました。特に、柴田克成氏には、市瀬夏洋氏(現東大計数工学科)、甲元洋氏(日立)と共に研究内容についてもしばしば議論の相手をして頂きました。ここに、感謝の意を表させていただきます。

岡部研究室の五月女悦久氏には、研究室の計算機管理の面で、頻繁にお世話になりました。また、氏の博学は、まさに驚嘆に値するものであり、私自身、自らの浅学を身にしみること度々でありました。この三年間に彼から得た知識の量は、私が彼に出会うまでの人生で得た知識の蓄積を遥かに凌駕すると言っても過言ではありません。計算機管理においては、廣瀬研究室の杉山和之氏にも度々サポートを賜りました。研究面では、林秀樹氏(東芝)の博士論文は本研究を進める上での参考になりました。また、卒論生の井上浩一氏には一年間本研究の進行に大きな貢献して頂きました。更に、松浦幹太氏(現東大生産研)からはプレゼンテーションのコツ等を学びました。他にも、岡部研究室、廣瀬研究室の皆様には、事あるごとに暖かいご支援を頂きました。ありがとうございました。

本研究を二年間、資金面から支援して頂いた日本学術振興会に感謝致します。このサポートによって、経済的支障無く研究活動を進めることができました。

修士二年からの約四年間、数多くの学会、研究会において、自らの研究を発表する場を与えられ、それによって自らの研究成果を整理するとともに、新たな目標に向かってのモチベーションを保つことができました。これも、日本神経回路学会、電子情報通信学会ニューロコンピューティング研究会をはじめとする、ニューロ関係の学会、研究会の運営にご尽力されている多くの方々に支えられてのことだと思えます。ここにお礼を申し上げます。

また、博士課程の三年間は、自らの専門の枠を超えて、多くのことを学ぶ機会に恵まれたと思えます。日本数理科学振興会の主催するJAMSセミナーにおける多くの先生、学生、若手研究者との出会いは、自らの研究に対する刺激になると同時に、研究を超えて、文化的に、社会的に、多くの貴重な経験を私に与えてくれました。そのような場の提供にご尽力頂いた広中平祐先生(山口大学学長)、Daniel Kleppner先生(MIT教授)、Beatrice Kleppner先生をはじめとする多くの皆様に感謝の意を表したいと存じます。また、JAMSセミナー以外にも、世界エネルギー会議をはじめとする多くの議論の機会にも恵まれました。そこでの経験も、私の人生でのかけがえのない財産になっています。そのような場で私をご指導頂いた諸先生方、ご支援頂いた電気学会他の皆様にもこの場を借りてお礼申し上げます。

最後に、苦難において私を励ましてくれた多くの友人達、そして博士課程進学のモラトリアムを快く支援してくれた家族へ、心からの感謝の気持ちを最後の一文に込め、本論文の謝辞に代えさせて頂きたいと思えます。

参考文献

- [1] Aiyer, S. V. B., Niranjana, M., & Fallside, F. (1990): A theoretical investigation into the performance of the Hopfield model, *IEEE trans.*, **NN-1**, 204-215.
- [2] Amari, S. (1977): Neural theory of association and concept-formation, *Bio. Cybernetics*, **26**, 175-185.
- [3] 甘利俊一 (1978): 神経回路網の数理, 産業図書.
- [4] Amari, S. & Maginu, K. (1988): Statistical neurodynamics of associative memory, *Neural Networks*, **1**, 63-73.
- [5] Amari, S. (1989): Characteristics of sparsely encoded associative memory, *Neural Networks*, **2**, 451-457.
- [6] Amit, D. J., Gutfreund, H., & Sompolinsky, H. (1985): Spin-glass model of neural networks, *Phys. Rev.*, **32**, 1007-1018.
- [7] Amit, D. J., Gutfreund, H., & Sompolinsky, H. (1985): Storing infinite numbers of patterns in a spin-glass model of neural networks, *Phys. Rev. Lett.*, **55**, 1530-1533.
- [8] Anderson, J. A. (1972): A simple neural network generating interactive memory, *Mathematical Biosciences*, **14**, 197-220.
- [9] Cortes, C., Krogh, A., & Hertz, J.A. (1987): Hierarchical associative networks, *J. Phys. A: Math. Gen.*, **20**, 4449-4455.
- [10] deSilva, C. J. S. & Attikiouzel, Y. (1992): Hopfield network as discrete dynamical systems, *Proc. of IJCNN '92*, **3**, 115-120.
- [11] 越前 功, 椎野正寿 (1996): 日本物理学会講演概要集, **3**, 630.
- [12] Fontanari, J. F. & Koberle, R. (1988): Retrieval via non-equilibrium states in neural networks. *J. Phys. A: Math. Gen.*, **21**, L667-L671.
- [13] Fu, A. M. N. (1994): Statistical analysis of an autoassociative memory network. *Neural Computation*, **6**, 837-841.
- [14] Gardner, E. (1988): The space of interactions in neural network models, *J. Phys. A: Math. Gen.*, **21**, 257-270.

- [15] Gutfreund, H. (1988): "Neural networks with hierarchically correlated patterns", *Physical Review A*, vol. 37, 570-577.
- [16] 平原誠, 岡夏樹, 金道敏樹 (1994): 階層的に相関を持つパターンを記憶するカスケード連想記憶モデル, 信学技法, **NC94-58**, 57-64.
- [17] Hopfield, J.J. (1982): Neural networks and physical systems with emergent collective computational abilities, *Proc. Natl. Acad. Sci. USA*, **79**, 2554-2558.
- [18] Hopfield, J. J. and Tank, D. W. (1985): Neural computation of decisions in optimization problems, *Biological Cybernetics*, **52**, 141-152.
- [19] 金道敏樹 (1994): 連想記憶のダイナミクスの解析, 信学技報, **NC94-1**, 1-8.
- [20] 金道敏樹 (1995): 連想記憶のダイナミクスの幾何学的解析, 信学論, **J78-D-II**, 1877-1889.
- [21] Kirkpatrick, S., Gelatt Jr., C. D. and Vecchi, M. P. (1983): Optimization by Simulated Annealing, *Science*, **220**, 671-680.
- [22] Konohen, T. (1972): Correlation matrix memories, *IEEE Trans.*, **C-21**, 353-359.
- [23] Krogh, A. & Hertz, J.A. (1988): Mean-field analysis of hierarchical associative networks with 'magnetisation', *J. Phys. A: Math. Gen.*, **21**, 2211-2224.
- [24] 三村和史, 岡田真人 (1995): 日本神経回路学会誌, **2**, 38-42.
- [25] McCulloch, W.S. & Pitts, W. (1943): A logical calculus of the ideas immanent in nervous activity, *Bull. Math. Biophys.*, **5**, 115-133.
- [26] 森永聡, 吉澤修治 (1994): 連想記憶のイジングモデルへの α オーダーパラメータの提案 — 自由エネルギーの計算 —, 信学技報, **NC93-96**, 119-125.
- [27] 森永聡, 原田大志. (1995) : 連想記憶ニューラルネットワークへの α パラメータの提案 — 平衡状態・ダイナミクスの解析 —, 信学論, **J78-D-II**, 954-961.
- [28] 森田昌彦 (1991): 神経回路モデルによる脳の記憶機構の研究, 東京大学工学部計数工学科博士論文.
- [29] Morita, M. (1993): Associative memory with nonmonotonic dynamics, *Neural Networks*, **6**, 115-126.
- [30] Morita, M. (1996). Memory and Learning of Sequential Patterns by Nonmonotone Neural Networks. *Neural Networks*, **9**, 1477-1489.

- [31] Nakano, K. (1972): Associatron – A model of associative memory, *IEEE Trans.*, **SMC-2**, 380-388.
- [32] Parga, N., and Virasoro, M. A. (1986): “The ultrametric organization of memories in a neural network”, *J. Physique*, vol. 47, 1857-1864.
- [33] Rammal, R., Toulouse, G., and Virasoro, M. A. (1986): “Ultrametricity for physicists”, *Rev. Mod. Phys.*, vol. 58, 765-786.
- [34] Rumelhart, D. E., Hinton, G. E., and Williams. R. J. (1986): Learning representations by back-propagating errors, *Nature*, **323**, 9, 533-536.
- [35] Okada, M. (1995): A hierarchy of macrodynamical equations for associative memory. *Neural Networks*, **8**, 833-838.
- [36] 岡田真人 (1995): 日本神経回路学会第6回全国大会講演論文集, 38-39.
- [37] Parga, N. & Virasoro, M. A. (1986): The ultrametric organization of memories in a neural network, *J. Physique*, **47**, 1857-1864.
- [38] Rosch, E. (1973): Natural Categories, *Cognitive Psychology*, **4**, 328-350.
- [39] 重光 卓, 岡田真人 (1995): スパースコーディングされた自己相関型連想記憶モデルの引き込み領域, 信学技報, **NC94**-119, 41-48.
- [40] Shiino, M. & Fukai, T. (1992): Self-consistent signal-to-noise analysis and its application to analogue neural networks with asymmetric connections. *Journal of Physics*, **A-25**, L375-L381.
- [41] Shiino, M. & Fukai, T. (1993): Self-consistent signal-to-noise analysis of the statistical behavior of analog neural networks and enhancement of the storage capacity. *Phys. Rev.*, bf E-48, 867-897.
- [42] Yanai, H. & Sawada, Y. (1990): Associative memory network composed of neurons with hysteretic property, *Neural Networks*, **3**, 223-228.
- [43] Yanai, H. & Amari, S. (1996): Auto-associative memory with two-stage dynamics of nonmonotonic neurons. *IEEE Trans.*, **NN-7**, 803-815.
- [44] Yoshizawa, S., Morita, M., & Amari, S. (1993): Capacity of associative memory using a nonmonotonic neuron model, *Neural Netowrks*, **6**, 167-176.

付録 A : 離散時間対称ネットワークのエネルギー

< 定理1 > 「離散時間同期型ニューラルネットワークにおいて、各ニューロンが符号関数

$$x_i(t+1) = \text{sgn}(u_i(t)) \quad (\text{A.1})$$

を活性化関数として持ち、荷重結合が対称の場合、 $E(t) = -\mathbf{x}(t+1)^T W \mathbf{x}(t)$ がリアップノフ関数になっている (単調非増加である)。

$E(t) = -\mathbf{x}(t+1)^T W \mathbf{x}(t)$ は明らかに有界であり、 $E(t)$ が単調非増加ならば $E(t)$ はある値に収束する。 $E(t)$ が単調非増加かどうかを調べるために $\Delta E(t) = E(t+2) - E(t)$ を考える。 $\delta \mathbf{x}(t) = \mathbf{x}(t+2) - \mathbf{x}(t)$ とおくと、

$$\Delta E(t) = E(t+2) - E(t) \quad (\text{A.2})$$

$$= -\mathbf{x}(t+3)^T W \mathbf{x}(t+2) + \mathbf{x}(t+1)^T W \mathbf{x}(t) \quad (\text{A.3})$$

$$= -\left(\mathbf{x}(t+3)^T W \mathbf{x}(t+2) - \mathbf{x}(t+1)^T W \mathbf{x}(t+2)\right) - \left(\mathbf{x}(t+1)^T W \mathbf{x}(t+2) - \mathbf{x}(t+1)^T W \mathbf{x}(t)\right) \quad (\text{A.4})$$

$$= -\left(\mathbf{x}(t+3)^T W \mathbf{x}(t+2) - \mathbf{x}(t+1)^T W \mathbf{x}(t+2)\right) - \left(\mathbf{x}(t+2)^T W \mathbf{x}(t+1) - \mathbf{x}(t)^T W \mathbf{x}(t+1)\right) \quad (\text{A.5})$$

$$= -\delta \mathbf{x}(t+1)^T W \mathbf{x}(t+2) - \delta \mathbf{x}(t)^T W \mathbf{x}(t+1) \quad (\text{A.6})$$

となる。ここで式 (A.4) から式 (A.5) への変形で W が対称であることを利用した。

今、 $\sum_j w_{ij} x_j(t+2) \geq 0$ とすると、式 (A.1) より $x_i(t+3) = 1$ となり、 $\delta x_i(t+1) = x_i(t+3) - x_i(t+1)$ は2か0になるから、 $\delta x_i(t+1) \geq 0$ が成立する。また、 $\sum_j w_{ij} x_j(t+2) < 0$ とすると、 $x_i(t+3) = -1$ となり、今度は $\delta x_i(t+1) \leq 0$ が成立する。よって、

$$\delta \mathbf{x}(t+1)^T W \mathbf{x}(t+2) \geq 0 \quad (\text{A.7})$$

が導かれる。同様に

$$\delta \mathbf{x}(t)^T W \mathbf{x}(t+1) \geq 0 \quad (\text{A.8})$$

も成り立つ。以上のことから

$$\Delta E(t) = -\delta \mathbf{x}(t+1)^T W \mathbf{x}(t+2) - \delta \mathbf{x}(t)^T W \mathbf{x}(t+1) \leq 0 \quad (\text{A.9})$$

が得られ、 $E(t)$ がリアップノフ関数になっている (単調非増加である) ことが言える。

< 定理2 > 各ニューロンが符号関数の活性化関数を持ち、離散時間同期型ニューラルネットワークが対称結合を持つ場合、内部状態が0にならなければ、そのネットワークは静止状態か周期2の振動状態に落ち着く。

定理1で、リアプノフ関数 $E(t)$ から状態 $\mathbf{x}(t)$ が一意に定まれば、系は2周期以下の状態に必ず収束すると言える。しかし、ここではそれは必ず成り立つとは言えない。そこで、エネルギー変化が行われない時は状態が遷移しないことを示せば良い。

リアプノフ関数(エネルギー)に変化がないのは、式(A.6)から分かるように $\delta x_i(t), \delta x_i(t+1)$ が全ての i について0の場合か、 $\delta x_i(t)$ が0でない i について $\sum_j w_{ij} x_j(t+1) = 0$ 、 $\delta x_i(t+1)$ が0でない i について $\sum_j w_{ij} x_j(t+2) = 0$ となる場合に限られる。ここで、内部状態は0でないと仮定したから、 $\forall i \delta x_i(t) = \delta x_i(t+1) = 0$ が成り立つ。よって、時刻 t と時刻 $t+2$ の状態は完全に等しく、周期2以下の状態に収束することになる。

尚、結合 $W = [w_{ij}]$ として対称性の制約のもとランダムな実数を与えた場合、丁度 $\sum_j w_{ij} x_j(t) = 0$ となる確率は測度0であり、ここで仮定した内部状態が0に等しくないという命題は一般に成り立つものである。

付録 B : ニューロウインドウ法の SCSNA

N 個のニューロンからなるネットワークに、相関をもつパターンを自己相関学習で記録し、それを部分反転法で想起する場合の系の平衡状態のセルフコンシステントな方程式を求める。ここで、相関をもつ記憶パターンとしては、 aN 個の親パターン ξ^μ をランダムに生成し、それぞれの親パターンから相関 r をもつ子パターン $\xi^{(\mu,\nu)}$ を M 個ずつ作り、その子パターンを記憶パターンとする。

まず、部分反転法における平衡状態の式は

$$x_i^{(1)} = \phi\left(\sum_j w_{ij} x_j^{(2)}\right) \quad (\text{B.1})$$

$$x_i^{(2)} = f\left(\sum_j w_{ij} (x_j^{(2)} - x_j^{(1)})\right) \quad (\text{B.2})$$

で与えられる。

これを状態ベクトル \mathbf{x} と記憶パターン ξ の相関 m を使った式に書き換える。そのために、最初に想起されていないパターンとの相関を求める。以下では、まず、想起されていない子パターン $\xi^{(\mu,\nu)}$ に着目する。記述の簡単のため、注目するパターンを ξ^0 と表し、この着目した子パターンと相関があるパターンを ξ^a ($a \in A$)、相関がないパターンを ξ^b ($b \in B$) と記述する。

まず、記憶行列は

$$w_{ij} = \frac{1}{N} \xi_i^0 \xi_j^0 + \frac{1}{N} \sum_A \xi_i^a \xi_j^a + \frac{1}{N} \sum_B \xi_i^b \xi_j^b \quad (\text{B.3})$$

と表される。また、パターン ξ^0 と状態 $\mathbf{x}^{(\tau)}$ の相関は

$$m_0^{(1)} = \frac{1}{N} \sum_i \xi_i^0 x_i^{(1)} \quad (\text{B.4})$$

$$m_0^{(2)} = \frac{1}{N} \sum_i \xi_i^0 x_i^{(2)} \quad (\text{B.5})$$

である。これらの式を使うと、相関による平衡状態の記述

$$m_0^{(1)} = \frac{1}{N} \sum_i \xi_i^0 \phi\left[\xi_i^0 m_0^{(2)} + \sum_A \xi_i^a m_a^{(2)} + \sum_B \xi_i^b m_b^{(2)}\right] \quad (\text{B.6})$$

$$\begin{aligned} m_0^{(2)} &= \frac{1}{N} \sum_i \xi_i^0 f\left[\xi_i^0 (m_0^{(2)} - m_0^{(1)}) + \sum_A \xi_i^a (m_a^{(2)} - m_a^{(1)}) \right. \\ &\quad \left. + \sum_B \xi_i^b (m_b^{(2)} - m_b^{(1)})\right] \end{aligned} \quad (\text{B.7})$$

が得られる。それぞれの式をテーラー展開すると、

$$m_0^{(1)} = \frac{1}{N} \sum_i \xi_i^0 \phi + \frac{1}{N} \sum_i \xi_i^0 \left(\xi_i^0 m_0^{(2)} + \sum_A \xi_i^a m_a^{(2)} \right) \phi' \quad (\text{B.8})$$

$$m_0^{(2)} = \frac{1}{N} \sum_i \xi_i^0 f + \frac{1}{N} \sum_i \xi_i^0 \left(\xi_i^0 (m_0^{(2)} - m_0^{(1)}) + \frac{1}{N} \sum_A \xi_i^a (m_a^{(2)} - m_a^{(1)}) \right) f' \quad (\text{B.9})$$

となる。更に、 $\frac{1}{N} \sum_i \xi_i^0 \xi_i^0 = 1$, $\frac{1}{N} \sum_i \xi_i^0 \xi_i^a = r^2$ の関係を使って式を整理すると、

$$m_0^{(1)} = \frac{1}{N} \sum_i \xi_i^0 \phi + m_0^{(2)} \phi' + r^2 \sum_A m_a^{(2)} \phi' \quad (\text{B.10})$$

$$m_0^{(2)} = \frac{1}{N} \sum_i \xi_i^0 f + (m_0^{(2)} - m_0^{(1)}) f' + r^2 \sum_A (m_a^{(2)} - m_a^{(1)}) f' \quad (\text{B.11})$$

となる。

さて、ここで注目の対象を一つのパターンから、グループ A のパターン全体に拡張する。グループ A の 1 番目から M 番目までのパターンとの相関を m_1, m_2, \dots, m_M と記述することにする。そして、 $\mathbf{m}^{(1)} = (m_1^{(1)}, m_2^{(1)}, \dots, m_M^{(1)})^T$, $\mathbf{m}^{(2)} = (m_1^{(2)}, m_2^{(2)}, \dots, m_M^{(2)})^T$ と置く。また、 $\frac{1}{N} \sum_i \xi_i^a \phi_i = \bar{\phi}^a$, $\frac{1}{N} \sum_i \xi_i^a f_i = \bar{f}^a$ と簡略表示し、 $\boldsymbol{\phi} = (\bar{\phi}^1, \bar{\phi}^2, \dots, \bar{\phi}^M)^T$, $\mathbf{f} = (\bar{f}^1, \bar{f}^2, \dots, \bar{f}^M)^T$ とする。この時、上記の式は

$$\begin{pmatrix} \mathbf{m}^{(1)} \\ \mathbf{m}^{(2)} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\phi} \\ \mathbf{f} \end{pmatrix} + \begin{pmatrix} O & \Phi \\ -F & F \end{pmatrix} \begin{pmatrix} \mathbf{m}^{(1)} \\ \mathbf{m}^{(2)} \end{pmatrix} \quad (\text{B.12})$$

と書ける。ここで

$$\Phi_{ij} = \phi' \delta_{ij} + r^2 \phi' (1 - \delta_{ij}) \quad (\text{B.13})$$

$$F_{ij} = f' \delta_{ij} + r^2 f' (1 - \delta_{ij}) \quad (\text{B.14})$$

である。よって、 $\begin{pmatrix} O & \Phi \\ -F & F \end{pmatrix} = H$ と置き、 $(I - H)^{-1} = K = \begin{pmatrix} K^{(1,1)} & K^{(1,2)} \\ K^{(2,1)} & K^{(2,2)} \end{pmatrix}$ と置けば、

$$\mathbf{m}^{(1)} = K^{(1,1)} \boldsymbol{\phi} + K^{(1,2)} \mathbf{f} \quad (\text{B.15})$$

$$\mathbf{m}^{(2)} = K^{(2,1)} \boldsymbol{\phi} + K^{(2,2)} \mathbf{f} \quad (\text{B.16})$$

のように相関 m が計算される。

このように算出された m を使って、ニューロンの内部ポテンシャル h_i を求める。内部ポテンシャルは、関数 ϕ に関する内部ポテンシャル $h_i^{(1)} = \sum_j w_{ij} x_j^{(1)}$ と、関数 f に関する内部ポテンシャル $h_i^{(2)} = \sum_j w_{ij} (x_j^{(2)} - x_j^{(1)})$ があり、それぞれ $h_i^{(1)} = \sum_\nu \xi_i^\nu m_\nu^{(2)}$, $h_i^{(2)} = \sum_\nu \xi_i^\nu (m_\nu^{(2)} - m_\nu^{(1)})$ と書ける。ここでは、まず $h_i^{(1)}$ について考える。 $h_i^{(2)}$ については、後段で $h_i^{(1)}$ の場合と全く同様に計算できることを示す。

今、想起しているパターンを $\xi^{(1,1)}$ とする。パターン $\xi^{(\mu,\nu)}$ との相関を $m_{(\mu,\nu)}^{(2)}$ と記述すると、内部ポテンシャル $h_i^{(1)}$ は、

$$h_i^{(1)} = \sum_{\nu} \xi_i^{(1,\nu)} m_{(1,\nu)}^{(2)} + \sum_{\mu \neq 1} \sum_{\nu} \xi_i^{(\mu,\nu)} m_{(\mu,\nu)}^{(2)} \quad (\text{B.17})$$

となる。この右辺の第一項がシグナル項、第二項がノイズ項である。このノイズ項を求めるのが以下の課題となる。ノイズ項は異なる μ に関しても同じでかつ互いに無相関であるから、後で簡単に足し合わせることができる。よって、 $N_i^{(1)} = \sum_{\nu} \xi^{(\mu,\nu)} m_{(\mu,\nu)}^{(2)}$ を考えれば良い。これを具体的に計算すると、

$$E[N_i^{(1)}] = E\left[\sum_{\nu} \xi_i^{\nu} \frac{1}{N} \left[\sum_{\nu} \sum_{\lambda} K_{\nu\lambda}^{(2,1)} \sum_j \xi_j^{\lambda} \phi_j + \sum_{\nu} \sum_{\lambda} K_{\nu\lambda}^{(2,2)} \sum_j \xi_j^{\lambda} f_j \right]\right] \quad (\text{B.18})$$

$$\begin{aligned} &= \frac{1}{N} E\left[\sum_{\nu} K_{\nu\nu}^{(2,1)} \phi_i + \sum_{\nu} K_{\nu\nu}^{(2,2)} f_i + r^2 \sum_{\nu} \sum_{\lambda \neq \nu} K_{\nu\lambda}^{(2,1)} \phi_i + r^2 \sum_{\nu} \sum_{\lambda \neq \nu} K_{\nu\lambda}^{(2,2)} f_i\right] \\ &+ \frac{1}{N} E\left[\sum_{\nu} \sum_{\lambda} K_{\nu\lambda}^{(2,1)} \sum_{j \neq i} \xi_i^{\nu} \xi_j^{\lambda} \phi_j + \sum_{\nu} \sum_{\lambda} K_{\nu\lambda}^{(2,2)} \sum_{j \neq i} \xi_i^{\nu} \xi_j^{\lambda} f_j\right] \end{aligned} \quad (\text{B.19})$$

となる。ここで、

$$N_i^{(1)} = G_i^{(1)} + Z_i^{(1)} \quad (\text{B.20})$$

$$G_i^{(1)} = \frac{1}{N} E\left[\sum_{\nu} K_{\nu\nu}^{(2,1)} \phi_i + \sum_{\nu} K_{\nu\nu}^{(2,2)} f_i + r^2 \sum_{\nu} \sum_{\lambda \neq \nu} K_{\nu\lambda}^{(2,1)} \phi_i + r^2 \sum_{\nu} \sum_{\lambda \neq \nu} K_{\nu\lambda}^{(2,2)} f_i\right] \quad (\text{B.21})$$

$$Z_i^{(1)} = \frac{1}{N} E\left[\sum_{\nu} \sum_{\lambda} K_{\nu\lambda}^{(2,1)} \sum_{j \neq i} \xi_i^{\nu} \xi_j^{\lambda} \phi_j + \sum_{\nu} \sum_{\lambda} K_{\nu\lambda}^{(2,2)} \sum_{j \neq i} \xi_i^{\nu} \xi_j^{\lambda} f_j\right] \quad (\text{B.22})$$

と表記することにする。このうち Z_i は、 $\xi_i^{\nu}, \xi_j^{\lambda}$ が無相関であると考え、正規分布に従うとする。その分散 $\sigma_1^2 = E[Z_i^2]$ を、

$$\sigma_1^2 = E\left[\left(\frac{1}{N} \sum_{\nu} \sum_{\lambda} [K_{\nu\lambda}^{(2,1)} \sum_{j \neq i} \xi_i^{\nu} \xi_j^{\lambda} \phi_j + K_{\nu\lambda}^{(2,2)} \sum_{j \neq i} \xi_i^{\nu} \xi_j^{\lambda} f_j]\right)^2\right] \quad (\text{B.23})$$

から求める。今、 $E[\sum_{j \neq i} \phi_i \phi_j] = 0, E[\sum_{j \neq i} \phi_i f_j] = 0, E[\sum_{j \neq i} f_i f_j] = 0$ を仮定し、また表記の簡単のため $\phi = f^{(1)}, f = f^{(2)}$ とおくと、分散は

$$\begin{aligned} \sigma_1^2 &= E\left[\frac{1}{N^2} \sum_{\nu} \sum_{\lambda} \sum_{\nu'} \sum_{\lambda'} \left(\sum_{\tau=1,2} \sum_{\tau'=1,2} [K_{\nu\lambda}^{(2,\tau)} K_{\nu'\lambda'}^{(2,\tau')} \sum_{j \neq i} \xi_i^{\nu} \xi_j^{\lambda} f_j^{(\tau)}] \right. \right. \\ &\quad \left. \left. \times [K_{\nu\lambda}^{(2,\tau)} K_{\nu'\lambda'}^{(2,\tau')} \sum_{j \neq i} \xi_i^{\nu'} \xi_j^{\lambda'} f_j^{(\tau')}] \right)\right] \end{aligned} \quad (\text{B.24})$$

$$\begin{aligned} &= \frac{1}{N^2} E\left[\sum_{\nu} \sum_{\lambda} \sum_{\tau=1,2} \sum_{\tau'=1,2} K_{\nu\lambda}^{(2,\tau)} K_{\nu\lambda}^{(2,\tau')} f_j^{(\tau)} f_j^{(\tau')} \right. \\ &\quad \left. + r^2 \sum_{\nu} \sum_{\lambda} \sum_{\nu' \neq \nu} \sum_j K_{\nu\lambda}^{(2,\tau)} K_{\nu'\lambda}^{(2,\tau')} f_j^{(\tau)} f_j^{(\tau')} \right] \end{aligned}$$

$$\begin{aligned}
 & +r^2 \sum_{\nu} \sum_{\lambda} \sum_{\lambda' \neq \lambda} \sum_j K_{\nu\lambda}^{(2,\tau)} K_{\nu\lambda'}^{(2,\tau')} f_j^{(\tau)} f_j^{(\tau')} \\
 & +r^4 \sum_{\nu} \sum_{\lambda} \sum_{\nu' \neq \nu} \sum_{\lambda' \neq \lambda} \sum_j K_{\nu\lambda}^{(2,\tau)} K_{\nu'\lambda'}^{(2,\tau')} f_j^{(\tau)} f_j^{(\tau')} \Big] \quad (\text{B.25})
 \end{aligned}$$

と計算される。 $Z_i^{(1)}$ の項は、この分散 σ_1^2 と確率変数 $z \sim N(0, 1)$ を使って

$$Z_i^{(1)} \sim \sigma_1 z \quad (\text{B.26})$$

と記述される。

内部ポテンシャル $h_i^{(2)} = \sum_j w_{ij}(x_j^{(2)} - x_j^{(1)})$ の場合は、

$$h_i^{(2)} = \sum_{\nu} \xi^{(1,\nu)} (m_{(1,\nu)}^{(2)} - m_{(1,\nu)}^{(1)}) + \sum_{\mu \neq 1} \sum_{\nu} \xi^{(\mu,\nu)} (m_{(\mu,\nu)}^{(2)} - m_{(\mu,\nu)}^{(1)}) \quad (\text{B.27})$$

となるので、 $h_i^{(1)}$ の場合と全く同じ式展開で、単に $K_{ij}^{(2,\lambda)}$ の代わりに $K_{ij}^{(2,\lambda)} - K_{ij}^{(1,\lambda)}$ を代入すればよい。

さて、ここまでで内部ポテンシャルを相関 m を使って記述することに成功したので、 m に関するセルフコンシステントな式

$$m_{(1,\mu)}^{(1)} = \sum_i \xi_i^{(1,\mu)} \phi \left(\sum_{\nu} \xi^{(1,\nu)} m_{(1,\nu)}^{(2)} \right) + aNG_i^{(1)} + aNZ_i^{(1)} \quad (\text{B.28})$$

$$m_{(1,\mu)}^{(2)} = \sum_i \xi_i^{(1,\mu)} f \left(\sum_{\nu} \xi^{(1,\nu)} (m_{(1,\nu)}^{(2)} - m_{(1,\nu)}^{(1)}) \right) + aNG_i^{(2)} + aNZ_i^{(2)} \quad (\text{B.29})$$

を得ることができる。ここで、

$$G_i^{(\rho)} = \frac{1}{N} E \left[\sum_{\tau=1,2} \sum_{\nu} V_{\nu\nu}^{(\rho,\tau)} f_i^{(\tau)} + r^2 \sum_{\tau=1,2} \sum_{\nu} \sum_{\lambda \neq \nu} V_{\nu\lambda}^{(\rho,\tau)} f_i^{(\tau)} \right] \quad (\text{B.30})$$

$$Z_i^{(\rho)} \sim \sigma_{\rho} z \quad (\text{B.31})$$

$$\begin{aligned}
 \sigma_{\rho}^2 & = \frac{1}{N^2} E \left[\sum_{\nu} \sum_{\lambda} \sum_{\tau=1,2} \sum_{\tau'=1,2} V_{\nu\lambda}^{(\rho,\tau)} V_{\nu\lambda}^{(\rho,\tau')} f_j^{(\tau)} f_j^{(\tau')} \right. \\
 & + r^2 \sum_{\nu} \sum_{\lambda} \sum_{\nu' \neq \nu} \sum_j V_{\nu\lambda}^{(\rho,\tau)} V_{\nu'\lambda}^{(\rho,\tau')} f_j^{(\tau)} f_j^{(\tau')} \\
 & + r^2 \sum_{\nu} \sum_{\lambda} \sum_{\lambda' \neq \lambda} \sum_j V_{\nu\lambda}^{(\rho,\tau)} V_{\nu\lambda'}^{(\rho,\tau')} f_j^{(\tau)} f_j^{(\tau')} \\
 & \left. + r^4 \sum_{\nu} \sum_{\lambda} \sum_{\nu' \neq \nu} \sum_{\lambda' \neq \lambda} \sum_j V_{\nu\lambda}^{(\rho,\tau)} V_{\nu'\lambda'}^{(\rho,\tau')} f_j^{(\tau)} f_j^{(\tau')} \right] \quad (\text{B.32})
 \end{aligned}$$

$$z \sim N(0, 1) \quad (\text{B.33})$$

である。但し、 $V^{(1,\tau)} = K^{(2,\tau)}$, $V^{(2,\tau)} = K^{(2,\tau)} - K^{(1,\tau)}$ と置いた。各平均量を

$$q_{11} = \frac{1}{N} \sum_i \phi_i^2 \quad (\text{B.34})$$

$$q_{12} = q_{21} = \frac{1}{N} \sum_i \phi_i f_i \quad (\text{B.35})$$

$$q_{22} = \frac{1}{N} \sum_i f_i^2 \quad (\text{B.36})$$

$$U_1 = \frac{1}{N} \sum_i \phi'_i \quad (\text{B.37})$$

$$U_2 = \frac{1}{N} \sum_i f'_i \quad (\text{B.38})$$

と記述し、以上の計算を総合すると、最終的にセルフコンシステントな関係式 (6.17)~(6.30) が求まる。

付録 C： 部分反転法の統計神経力学

統計神経力学では、各ニューロンへの入力をシグナルとノイズに分割する。そして、ノイズ項を正規分布と仮定し、その分散の変化とシグナル項の変化を漸化式の形に持ち込む。ここでは、まず一次近似について考える。

今、記憶した aN 個のパターンのうち 1 番目のパターンを想起する過程を考える。すると、各ニューロンの状態遷移は、部分反転法の第一ステップでは

$$x_i^{t+1} = f_1[\xi_i^1 m_t^1 + \frac{1}{N} \sum_{j \neq i} \sum_{\mu=2}^{aN} \xi_i^\mu \xi_j^\mu x_j^t] \quad (\text{C.1})$$

第二ステップでは

$$x_i^{t+2} = f_0[\xi_i^1 (m_t^1 - m_{t+1}^1) + \frac{1}{N} \sum_{j \neq i} \sum_{\mu=2}^{aN} \xi_i^\mu \xi_j^\mu (x_j^t - x_j^{t+1})] \quad (\text{C.2})$$

となる。ここで、 $x^t (t: \text{偶数})$ が観測される状態ベクトルとなる。

以下、ノイズ項を

$$z_i^t = \frac{1}{N} \sum_{j \neq i} \sum_{\mu=2}^{aN} \xi_i^\mu \xi_j^\mu x_j^t \quad (\text{C.3})$$

$$z_i^{t+1} = \frac{1}{N} \sum_{j \neq i} \sum_{\mu=2}^{aN} \xi_i^\mu \xi_j^\mu x_j^{t+1} \quad (\text{C.4})$$

$$z_i^{t+2} = \frac{1}{N} \sum_{j \neq i} \sum_{\mu=2}^{aN} \xi_i^\mu \xi_j^\mu x_j^{t+2} \quad (\text{C.5})$$

と表記する。ここでは、このノイズ項が正規分布であると仮定して話を進めるので、 z_i^t の分散 σ_t がどのような漸化式で表されるかを求めるのが目標になる。

ノイズ項は、 x がランダムな変数であれば、その分散は簡単に求まる。しかし、実際は、これらは行列 $\sum_{\mu} \xi_i^\mu \xi_j^\mu$ による変換をうけてきた経歴をもつため、 $\xi_i^\mu \xi_j^\mu$ と相関をもつことになる。 x を $\xi_i^\mu \xi_j^\mu$ との相関をもつ項ともたない項に分けるため、テーラー展開を用いると

$$z_i^{t+1} = \frac{1}{N} \sum_{j \neq i} \sum_{\mu=2}^{aN} \xi_i^\mu \xi_j^\mu x_j^{t+1,(\mu)} + U_{t+1} z_i^t + x_i^t a U_{t+1} \quad (\text{C.6})$$

$$x_j^{t+1,(\mu)} = f_1[\xi_j^1 m_t^1 + \frac{1}{N} \sum_{k \neq j} \sum_{\nu \neq 1, \mu}^{aN} \xi_j^\nu \xi_k^\nu x_k^t] \quad (\text{C.7})$$

$$U_{t+1} = \frac{1}{N} \sum_{j=1}^N f_1'[\xi_j^1 m_t^1 + \frac{1}{N} \sum_{k \neq j} \sum_{\nu=2}^{aN} \xi_j^\nu \xi_k^\nu x_k^t] \quad (\text{C.8})$$

$$z_i^{t+2} = \frac{1}{N} \sum_{j \neq i} \sum_{\mu=2}^{aN} \xi_i^\mu \xi_j^\mu x_j^{t+2,(\mu)} + U_{t+2}(z_i^t - z_i^{t+1}) + (x_i^t - x_i^{t+1})aU_{t+2} \quad (C.9)$$

$$x_j^{t+2,(\mu)} = f_0[\xi_j^1(m_t^1 - m_{t+1}^1) + \frac{1}{N} \sum_{k \neq j} \sum_{\nu \neq 1, \mu}^{aN} \xi_j^\nu \xi_k^\nu (x_k^t - x_k^{t+1})] \quad (C.10)$$

$$U_{t+2} = \frac{1}{N} \sum_{j=1}^N f_0'[\xi_j^1(m_t^1 - m_{t+1}^1) + \frac{1}{N} \sum_{k \neq j} \sum_{\nu=2}^{aN} \xi_j^\nu \xi_k^\nu (x_k^t - x_k^{t+1})] \quad (C.11)$$

が得られる。ここで、ノイズ項の平均は0であると仮定すると、ノイズ項は

$$z_i^{t+1} = \frac{1}{N} \sum_{j \neq i} \sum_{\mu=2}^{aN} \xi_i^\mu \xi_j^\mu x_j^{t+1,(\mu)} \quad (C.12)$$

$$z_i^{t+2} = \frac{1}{N} \sum_{j \neq i} \sum_{\mu=2}^{aN} \xi_i^\mu \xi_j^\mu x_j^{t+2,(\mu)} + U_{t+2}(z_i^t - z_i^{t+1}) \quad (C.13)$$

となる。これらの z_t, z_{t+1}, z_{t+2} をもとに、 $\rho_{t+1} = (E[(z_t - z_{t+1})^2])^{\frac{1}{2}}, \sigma_{t+2} = (E[z_{t+2}^2])^{\frac{1}{2}}$ を求めると、記憶パターンとのオーバーラップ m の漸化式

$$m_{t+1} = \int Dz \langle \xi f_1(\xi m_t + \sigma_t z) \rangle_\xi \quad (C.14)$$

$$m_{t+2} = \int Dz \langle \xi f_0(\xi(m_t - m_{t+1}) + \rho_{t+1} z) \rangle_\xi \quad (C.15)$$

$$Dz = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz \quad (C.16)$$

が得られる。よって、後はこの σ, ρ のダイナミクスが求まればよい。

σ と ρ を求めるには、 z の分散を計算することになるが、上で述べたように、これらは ξ と x の相関のため、過去の x に遡って計算をする必要がある。ここでは、 x_t 以降の相関を考え、 x_{t-1} 以前の ξ 依存性までは遡らないことにする。すると、 $z_i^t, z_i^{t+1}, z_i^{t+2}$ は

$$z_i^t = \frac{1}{N} \sum_{j \neq i} \sum_{\mu=2}^{aN} \xi_i^\mu \xi_j^\mu x_j^t \quad (C.17)$$

$$z_i^{t+1} = \frac{1}{N} \sum_{j \neq i} \sum_{\mu=2}^{aN} \xi_i^\mu \xi_j^\mu x_j^{t+1,(\mu)} + U_{t+1} \frac{1}{N} \sum_{j \neq i} \sum_{\mu=2}^{aN} \xi_i^\mu \xi_j^\mu x_j^t \quad (C.18)$$

$$z_i^{t+2} = \frac{1}{N} \sum_{j \neq i} \sum_{\mu=2}^{aN} \xi_i^\mu \xi_j^\mu x_j^{t+2,(\mu)} + U_{t+2}(z_i^t - z_i^{t+1}) \quad (C.19)$$

と展開すればよいことになる。 $\rho_{t+1} = (E[(z_t - z_{t+1})^2])^{\frac{1}{2}}, \sigma_{t+2} = (E[z_{t+2}^2])^{\frac{1}{2}}$ であるから、以上の z を使って、分散 $S_{t,t} = E[z_t, z_t], S_{t+1,t+1} = E[z_{t+1}^2], S_{t+2,t+2} = E[z_{t+2}^2]$ 、共分散 $S_{t,t+1} = E[z_t z_{t+1}]$ を計算すれば、 σ, ρ は求まる。具体的に S を計算すると

$$S_{t,t} = E\left[\left(\frac{1}{N} \sum_{j \neq i} \sum_{\mu=2}^{aN} \xi_i^\mu \xi_j^\mu x_j^{t+1,(\mu)}\right)^2\right] \quad (C.20)$$

$$= aQ_{t,t} \quad (C.21)$$

$$S_{t,t+1} = E\left[\left(\frac{1}{N} \sum_{j \neq i} \sum_{\mu=2}^{aN} \xi_i^\mu \xi_j^\mu x_j^t\right) \left(\frac{1}{N} \sum_{j \neq i} \sum_{\mu=2}^{aN} \xi_i^\mu \xi_j^\mu (x_j^{t+1,(\mu)} + U_{t+1} x_j^t)\right)\right] \quad (C.22)$$

$$= aQ_{t,t+1} + aU_{t+1}Q_{t,t} \quad (C.23)$$

$$= aQ_{t,t+1} + U_{t+1}S_{t,t} \quad (C.24)$$

$$S_{t+1,t+1} = E\left[\left(\frac{1}{N} \sum_{j \neq i} \sum_{\mu=2}^{aN} \xi_i^\mu \xi_j^\mu x_j^{t+1,(\mu)} + U_{t+1} z_i^t\right)^2\right] \quad (C.25)$$

$$= E\left[\left(\frac{1}{N} \sum_{j \neq i} \sum_{\mu=2}^{aN} \xi_i^\mu \xi_j^\mu (x_j^{t+1,(\mu)} + U_{t+1} x_j^t)\right)^2\right] \quad (C.26)$$

$$= aQ_{t+1,t+1} + aU_{t+1}^2 Q_{t,t} + 2aU_{t+1}Q_{t,t+1} \quad (C.27)$$

$$= aQ_{t+1,t+1} + U_{t+1}^2 S_{t,t} + 2aU_{t+1}Q_{t,t+1} \quad (C.28)$$

$$S_{t+2,t+2} = E\left[\left(\frac{1}{N} \sum_{j \neq i} \sum_{\mu=2}^{aN} \xi_i^\mu \xi_j^\mu x_j^{t+2,(\mu)} + U_{t+2}(z_i^t - z_i^{t+1})\right)^2\right] \quad (C.29)$$

$$= E\left[\left(\frac{1}{N} \sum_{j \neq i} \sum_{\mu=2}^{aN} \xi_i^\mu \xi_j^\mu (x_j^{t+2,(\mu)} + U_{t+2} x_j^t - U_{t+2}(x_j^{t+1,(\mu)} + U_{t+1} x_j^t))\right)^2\right] \quad (C.30)$$

$$= E\left[\left(\frac{1}{N} \sum_{j \neq i} \sum_{\mu=2}^{aN} \xi_i^\mu \xi_j^\mu (x_j^{t+2,(\mu)} - U_{t+2} x_j^{t+1,(\mu)} + U_{t+2}(1 - U_{t+1}) x_j^t)\right)^2\right] \quad (C.31)$$

$$= aQ_{t+2,t+2} + aU_{t+2}^2 Q_{t+1,t+1} + aU_{t+2}^2 (1 - U_{t+1})^2 Q_{t,t} - aU_{t+2} Q_{t+1,t+2} + aU_{t+2}(1 - U_{t+1}) Q_{t,t+2} - aU_{t+2}^2 (1 - U_{t+1}) Q_{t,t+1} \quad (C.32)$$

$$= aQ_{t+2,t+2} + aU_{t+2}^2 Q_{t+1,t+1} + U_{t+2}^2 (1 - U_{t+1})^2 S_{t,t} - aU_{t+2} Q_{t+1,t+2} + aU_{t+2}(1 - U_{t+1}) Q_{t,t+2} - aU_{t+2}^2 (1 - U_{t+1}) Q_{t,t+1} \quad (C.33)$$

となる。ここで、 $Q_{t,\tau} = E[x_t x_\tau]$ である。この S を使って

$$\rho_{t+1}^2 = S_{t,t} + S_{t+1,t+1} - 2S_{t,t+1} \quad (C.34)$$

$$\sigma_{t+2,t+2}^2 = S_{t+2,t+2} \quad (C.35)$$

とノイズ項の分散を計算することができる。

しかし、これらを実際求めるには、以上の計算結果で現れた U と Q を得なくてはならない。まず、 U は

$$U_{t+1} = \frac{1}{\sigma_t} \int Dz z \langle \xi f_1(\xi m_t + \sigma_t z) \rangle_\xi \quad (C.36)$$

$$U_{t+2} = \frac{1}{\sigma_s} \int Dz z \langle \xi f_0(\xi(m_t - m_{t+1}) + \rho_{t+1} z) \rangle_\xi \quad (C.37)$$

で求められる。

次に $Q_{t,t}$, $Q_{t+1,t+1}$, $Q_{t+2,t+2}$, $Q_{t,t+1}$, $Q_{t,t+2}$, $Q_{t+1,t+2}$ を求める。まず、 $t \neq \tau$ の場合を考える。 x の間で相関を計算する時、前の時刻における z の相関を考慮する必要がある。し

かし、 $Q_{t,t+1}, Q_{t,t+2}$ の場合は、 $t-1$ 以前の相関を時間は考えなくてよいので、 z の相関を考慮する必要はない。よって、問題になるのは $Q_{t+1,t+2}$ だけである。この計算を行うと

$$Q_{t+1,t+2} = \int Dz_a \int Dz_b \int Dz_c < f_1(\xi m_t + \sigma_t(\sigma_t^{(\alpha)} z_a + \sigma_t^{(\beta)} z_b) + \Gamma_t) \times f_0(\xi(m_t - m_{t+1}) + \text{sgn}(C_{t+1,t+2})\rho_{t+1}(\sigma_t^{(\alpha)} z_a + \sigma_t^{(\beta)} z_c) + \Gamma_{t+1}) >_\xi \quad (\text{C.38})$$

$$C_{t+1,t+2} = (1 - U_{t+1})S_{t,t} - aQ_{t,t+1} \quad (\text{C.39})$$

$$\sigma_t^{(\alpha)} = \sqrt{\frac{|C_{t+1,t+2}|}{\sigma_t \rho_{t+1}}} \quad (\text{C.40})$$

$$\sigma_t^{(\beta)} = \sqrt{1 - \frac{|C_{t+1,t+2}|}{\sigma_t \rho_{t+1}}} \quad (\text{C.41})$$

が導かれる。他の $Q_{t,\tau}$ は $Q_{t,\tau} = m_t^1 m_\tau^1$ と計算される。次に、 $t = \tau$ の場合だけ考える。 $Q_{t,t}$ は上の議論から $Q_{t,t} = \frac{1}{a} S_{t,t}$ で与えられ、 $Q_{t+1,t+1}, Q_{t+2,t+2}$ は

$$Q_{t+1,t+1} = \int dz < f_1(\xi m_t + \sigma_t z)^2 >_\xi \quad (\text{C.42})$$

$$Q_{t+2,t+2} = \int dz < f_1(\xi(m_t - m_{t+1}) + \rho_{t+1} z)^2 >_\xi \quad (\text{C.43})$$

のように求めることができる。この結果をまとめると統計神経力学のダイナミクスが得られる。

以上の議論では、ノイズ項の平均は 0 と仮定していたが、ノイズ項の平均も考慮した場合を考える。ノイズ項を平均まで考えると

$$z_i^{t+1} = \frac{1}{N} \sum_{j \neq i} \sum_{\mu=2}^{aN} \xi_i^\mu \xi_j^\mu x_j^{t+1,(\mu)} + U_{t+1} z_i^t + a x_i^t U_{t+1} \quad (\text{C.44})$$

$$z_i^{t+2} = \frac{1}{N} \sum_{j \neq i} \sum_{\mu=2}^{aN} \xi_i^\mu \xi_j^\mu x_j^{t+2,(\mu)} + U_{t+2}(z_i^t - z_i^{t+1}) + a(x_i^t - x_i^{t+1})U_{t+2} \quad (\text{C.45})$$

となる。本来は i によって x は異なるが、これを平均化し、全ての i について一様に m^1 で近似すれば、上の式は

$$z_i^{t+1} = \frac{1}{N} \sum_{j \neq i} \sum_{\mu=2}^{aN} \xi_i^\mu \xi_j^\mu x_j^{t+1,(\mu)} + U_{t+1} z_i^t + a m^1 U_{t+1} \quad (\text{C.46})$$

$$z_i^{t+2} = \frac{1}{N} \sum_{j \neq i} \sum_{\mu=2}^{aN} \xi_i^\mu \xi_j^\mu x_j^{t+2,(\mu)} + U_{t+2}(z_i^t - z_i^{t+1}) + a(m_t^1 - m_{t+1}^1)U_{t+2} \quad (\text{C.47})$$

となる。このノイズ項の平均

$$\Gamma_{t+1} = a m_t U_{t+1} + U_{t+1} \Gamma_t \quad (\text{C.48})$$

$$\Gamma_{t+2} = a(m_t - m_{t+1})U_{t+2} + U_{t+2}(\Gamma_t - \Gamma_{t+1}) \quad (\text{C.49})$$

を上の計算結果に導入すると、本文中の式 (6.31)~(6.48) を得ることができる。

尚、二次近似の場合、ノイズ項

$$z_i^t = \frac{1}{N} \sum_{j \neq i}^N \sum_{\mu=2}^{aN} \xi_i^\mu \xi_j^\mu x_j^t \quad (\text{C.50})$$

$$z_i^{t+1} = \frac{1}{N} \sum_{j \neq i}^N \sum_{\mu=2}^{aN} \xi_i^\mu \xi_j^\mu x_j^{t+1} \quad (\text{C.51})$$

$$z_i^{t+2} = \frac{1}{N} \sum_{j \neq i}^N \sum_{\mu=2}^{aN} \xi_i^\mu \xi_j^\mu x_j^{t+2} \quad (\text{C.52})$$

$$z_i^{t+3} = \frac{1}{N} \sum_{j \neq i}^N \sum_{\mu=2}^{aN} \xi_i^\mu \xi_j^\mu x_j^{t+3} \quad (\text{C.53})$$

$$z_i^{t+4} = \frac{1}{N} \sum_{j \neq i}^N \sum_{\mu=2}^{aN} \xi_i^\mu \xi_j^\mu x_j^{t+4} \quad (\text{C.54})$$

を考え、それを

$$z_i^{t+1} = \frac{1}{N} \sum_{j \neq i}^N \sum_{\mu=2}^{aN} \xi_i^\mu \xi_j^\mu x_j^{t+1,(\mu)} + U_{t+1} z_i^t \quad (\text{C.55})$$

$$x_j^{t+1,(\mu)} = f_1[\xi_j^1 m_t^1 + \frac{1}{N} \sum_{k \neq j}^N \sum_{\nu \neq 1, \mu}^{aN} \xi_j^\nu \xi_k^\nu x_k^t] \quad (\text{C.56})$$

$$U_{t+1} = \frac{1}{N} \sum_{j=1}^N f_1'[\xi_j^1 m_t^1 + \frac{1}{N} \sum_{k \neq j}^N \sum_{\nu=2}^{aN} \xi_j^\nu \xi_k^\nu x_k^t] \quad (\text{C.57})$$

$$z_i^{t+2} = \frac{1}{N} \sum_{j \neq i}^N \sum_{\mu=2}^{aN} \xi_i^\mu \xi_j^\mu x_j^{t+2,(\mu)} + U_{t+2}(z_i^t - z_i^{t+1}) \quad (\text{C.58})$$

$$x_j^{t+2,(\mu)} = f_0[\xi_j^1(m_t^1 - m_{t+1}^1) + \frac{1}{N} \sum_{k \neq j}^N \sum_{\nu \neq 1, \mu}^{aN} \xi_j^\nu \xi_k^\nu (x_k^t - x_k^{t+1})] \quad (\text{C.59})$$

$$U_{t+2} = \frac{1}{N} \sum_{j=1}^N f_0'[\xi_j^1(m_t^1 - m_{t+1}^1) + \frac{1}{N} \sum_{k \neq j}^N \sum_{\nu=2}^{aN} \xi_j^\nu \xi_k^\nu (x_k^t - x_k^{t+1})] \quad (\text{C.60})$$

$$z_i^{t+3} = \frac{1}{N} \sum_{j \neq i}^N \sum_{\mu=2}^{aN} \xi_i^\mu \xi_j^\mu x_j^{t+3,(\mu)} + U_{t+3} z_i^{t+2} \quad (\text{C.61})$$

$$x_j^{t+3,(\mu)} = f_1[\xi_j^1 m_{t+2}^1 + \frac{1}{N} \sum_{k \neq j}^N \sum_{\nu \neq 1, \mu}^{aN} \xi_j^\nu \xi_k^\nu x_k^{t+2}] \quad (\text{C.62})$$

$$U_{t+3} = \frac{1}{N} \sum_{j=1}^N f'_1[\xi_j^1 m_{t+2}^1 + \frac{1}{N} \sum_{k \neq j}^N \sum_{\nu=2}^{aN} \xi_j^\nu \xi_k^\nu x_k^{t+2}] \quad (\text{C.63})$$

$$z_i^{t+4} = \frac{1}{N} \sum_{j \neq i}^N \sum_{\mu=2}^{aN} \xi_i^\mu \xi_j^\mu x_j^{t+4,(\mu)} + U_{t+4}(z_i^{t+2} - z_i^{t+3}) \quad (\text{C.64})$$

$$x_j^{t+4,(\mu)} = f_0[\xi_j^1(m_{t+2}^1 - m_{t+3}^1) + \frac{1}{N} \sum_{k \neq j}^N \sum_{\nu \neq 1, \mu}^{aN} \xi_j^\nu \xi_k^\nu (x_k^{t+2} - x_k^{t+3})] \quad (\text{C.65})$$

$$U_{t+4} = \frac{1}{N} \sum_{j=1}^N f'_0[\xi_j^1(m_{t+2}^1 - m_{t+3}^1) + \frac{1}{N} \sum_{k \neq j}^N \sum_{\nu=2}^{aN} \xi_j^\nu \xi_k^\nu (x_k^{t+2} - x_k^{t+3})] \quad (\text{C.66})$$

のようにテーラー展開し、これに基づき一次近似と同様の計算を用いて漸化式を導くことになる。一般に、 n 次の近似では、 $z_t, z_{t+1}, \dots, z_{t+2n-1}$ をテーラー展開し、時刻 t 以降の相関を考慮して漸化式を導く。

付録 D : ニューロウインドウ法の統計神経力学

ニューロウインドウ連想記憶の場合、ダイナミクス自体は部分反転法と同じである。しかし、ニューロウインドウ連想記憶ではパターンに相関がある場合を想定しているのので、それを考慮に入れた式展開が必要になる。以下では、統計神経力学の一次近似を考える。

今、相関が二段で、各クラスタが互いの相関が r であるような M 個のパターンからなる場合を考える。各クラスタを $\Lambda_1, \Lambda_2, \dots, \Lambda_{aN}$ で表記することにする。記憶した aN 個のクラスタのうち Λ_1 番目のクラスタに属するパターン ξ^1 を想起する過程を考える。すると、各ニューロンの状態遷移は、部分反転法の第一ステップでは

$$x_i^{t+1} = f_1 \left[\sum_{\lambda \in \Lambda_1} \xi_i^\lambda m_t^\lambda + \frac{1}{N} \sum_{j \neq i} \sum_{p=2}^{aN} \sum_{\mu \in \Lambda_p} \xi_i^\mu \xi_j^\mu x_j^t \right] \quad (\text{D.1})$$

第二ステップでは

$$x_i^{t+2} = f_0 \left[\sum_{\lambda \in \Lambda_1} \xi_i^\lambda (m_t^\lambda - m_{t+1}^\lambda) + \frac{1}{N} \sum_{j \neq i} \sum_{p=2}^{aN} \sum_{\mu \in \Lambda_p} \xi_i^\mu \xi_j^\mu (x_j^t - x_j^{t+1}) \right] \quad (\text{D.2})$$

となる。以下、ノイズ項を

$$z_i^t = \frac{1}{N} \sum_{j \neq i} \sum_{p=2}^{aN} \sum_{\mu \in \Lambda_p} \xi_i^\mu \xi_j^\mu x_j^t \quad (\text{D.3})$$

$$z_i^{t+1} = \frac{1}{N} \sum_{j \neq i} \sum_{p=2}^{aN} \sum_{\mu \in \Lambda_p} \xi_i^\mu \xi_j^\mu x_j^{t+1} \quad (\text{D.4})$$

$$z_i^{t+2} = \frac{1}{N} \sum_{j \neq i} \sum_{p=2}^{aN} \sum_{\mu \in \Lambda_p} \xi_i^\mu \xi_j^\mu x_j^{t+2} \quad (\text{D.5})$$

と表記する。ここでは、このノイズ項が正規分布であると仮定して話を進めるので、これらの z_t, z_{t+1}, z_{t+2} をもとに、 $\rho_{t+1} = (E[(z_t - z_{t+1})^2])^{\frac{1}{2}}, \sigma_{t+2} = (E[z_{t+2}^2])^{\frac{1}{2}}$ を求めると、記憶パターンとのオーバーラップ m の漸化式

$$m_{t+1} = \int Dz \langle \xi f_1(\xi m_t + \sigma_t z) \rangle_\xi \quad (\text{D.6})$$

$$m_{t+2} = \int Dz \langle \xi f_0(\xi(m_t - m_{t+1}) + \rho_{t+1} z) \rangle_\xi \quad (\text{D.7})$$

$$Dz = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz \quad (\text{D.8})$$

が得られる。よって、後はこの σ, ρ のダイナミクスがどのような漸化式で表されるかを求めるのが目標になる。ところが、上記の形のままで σ, ρ を求めるのは、 ξ 間の相関があるため容易ではない。そこで、

$$\sum_{\mu \in A_p} \xi_i^\mu \xi_j^\mu = \sum_{\kappa \in A_p} \lambda_\kappa \xi_i^\kappa \xi_j^\kappa \quad (D.9)$$

$$\frac{1}{N} \sum_{i=1}^N \xi_i^\kappa \xi_i^\eta = 0 \quad (\kappa \neq \eta, \kappa, \eta \in A_p) \quad (D.10)$$

$$\frac{1}{N} \sum_{i=1}^N \xi_i^\kappa \xi_i^\kappa = 1 \quad (D.11)$$

と書き換えることにする。ここで、クラスター中の相関が r でパターン数が M の場合、 λ_κ は

$$k_{ij} = \begin{cases} 1, & \text{if } i = j \\ r^2, & \text{if } i \neq j \end{cases} \quad (D.12)$$

を満たす $M \times M$ 行列 $K = [k_{ij}]$ の固有値、 ξ^κ はその固有ベクトルとなっているので、それらを計算することで上記の表現を得ることができる。

さて、このように相関が除去されたノイズ項を用いると、 x は

$$x_i^{t+1} = f_1 \left[\sum_{\lambda \in \Lambda_1} \xi_i^\lambda m_t^\lambda + \frac{1}{N} \sum_{j \neq i} \sum_{p=2}^{aN} \sum_{\kappa \in A_p} \lambda_\kappa \xi_i^\kappa \xi_j^\kappa x_j^t \right] \quad (D.13)$$

$$x_i^{t+2} = f_0 \left[\sum_{\lambda \in \Lambda_1} \xi_i^\lambda (m_t^\lambda - m_{t+1}^\lambda) + \frac{1}{N} \sum_{j \neq i} \sum_{p=2}^{aN} \sum_{\kappa \in A_p} \lambda_\kappa \xi_i^\kappa \xi_j^\kappa (x_j^t - x_j^{t+1}) \right] \quad (D.14)$$

となり、 z は

$$z_i^t = \frac{1}{N} \sum_{j \neq i} \sum_{p=2}^{aN} \sum_{\kappa \in A_p} \lambda_\kappa \xi_i^\kappa \xi_j^\kappa x_j^t \quad (D.15)$$

$$z_i^{t+1} = \frac{1}{N} \sum_{j \neq i} \sum_{p=2}^{aN} \sum_{\kappa \in A_p} \lambda_\kappa \xi_i^\kappa \xi_j^\kappa x_j^{t+1} \quad (D.16)$$

$$z_i^{t+2} = \frac{1}{N} \sum_{j \neq i} \sum_{p=2}^{aN} \sum_{\kappa \in A_p} \lambda_\kappa \xi_i^\kappa \xi_j^\kappa x_j^{t+2} \quad (D.17)$$

と置き換えることができる。

ここで、 x の ξ 依存性を取り入れるため、 x_i^{t+1} と x_i^{t+2} をテーラー展開する。テーラー展開で出てくる 0 次、1 次微分の項をそれぞれ

$$x_i^{t+1,(p)} = f_1 \left[\sum_{\lambda \in \Lambda_1} \xi_i^\lambda m_t^\lambda + \frac{1}{N} \sum_{j \neq i} \sum_{q \neq 1,p} \sum_{\kappa \in A_q} \lambda_\kappa \xi_i^\kappa \xi_j^\kappa x_j^t \right] \quad (D.18)$$

$$U_{t+1} = \frac{1}{N} \sum_{j=1}^N f_1' \left[\sum_{\lambda \in \Lambda_1} \xi_i^\lambda m_t^\lambda + \frac{1}{N} \sum_{j \neq i} \sum_{p=2}^{aN} \sum_{\kappa \in A_p} \lambda_\kappa \xi_i^\kappa \xi_j^\kappa x_j^t \right] \quad (D.19)$$

$$x_i^{t+2,(p)} = f_0 \left[\sum_{\lambda \in \Lambda_1} \xi_i^\lambda (m_t^\lambda - m_{t+1}^\lambda) + \frac{1}{N} \sum_{j \neq i} \sum_{q \neq 1, p} \sum_{\kappa \in A_p} \lambda_\kappa \xi_i^\kappa \xi_j^\kappa (x_j^t - x_j^{t+1}) \right] \quad (\text{D.20})$$

$$U_{t+2} = \frac{1}{N} \sum_{j=1}^N f'_0 \left[\sum_{\lambda \in \Lambda_1} \xi_i^\lambda (m_t^\lambda - m_{t+1}^\lambda) + \frac{1}{N} \sum_{j \neq i} \sum_{p=2}^{aN} \sum_{\kappa \in A_p} \lambda_\kappa \xi_i^\kappa \xi_j^\kappa (x_j^t - x_j^{t+1}) \right] (\text{D.21})$$

となるので、 z_i^{t+1} , z_i^{t+2} はそれぞれ

$$\begin{aligned} z_i^{t+1} &= \frac{1}{N} \sum_{j \neq i} \sum_{p=2}^{aN} \sum_{\kappa \in A_p} \lambda_\kappa \xi_i^\kappa \xi_j^\kappa x_j^{t+1,(p)} \\ &\quad + \frac{1}{N^2} \sum_{j \neq i} \sum_{p=2}^{aN} \sum_{\kappa \in A_p} \lambda_\kappa \xi_i^\kappa \xi_j^\kappa \sum_{\eta \in A_p} \sum_{k=1}^N \lambda_\eta \xi_j^\eta \xi_k^\eta x_k^t U_{t+1} \end{aligned} \quad (\text{D.22})$$

$$= \frac{1}{N} \sum_{j \neq i} \sum_{p=2}^{aN} \sum_{\kappa \in A_p} \lambda_\kappa \xi_i^\kappa \xi_j^\kappa (x_j^{t+1,(p)} + \lambda_\kappa x_j^t U_{t+1}) \quad (\text{D.23})$$

$$\begin{aligned} z_i^{t+2} &= \frac{1}{N} \sum_{j \neq i} \sum_{p=2}^{aN} \sum_{\kappa \in A_p} \lambda_\kappa \xi_i^\kappa \xi_j^\kappa x_j^{t+2,(p)} \\ &\quad + \frac{1}{N^2} \sum_{j \neq i} \sum_{p=2}^{aN} \sum_{\kappa \in A_p} \lambda_\kappa \xi_i^\kappa \xi_j^\kappa \sum_{\eta \in A_p} \sum_{k=1}^N \lambda_\eta \xi_j^\eta \xi_k^\eta (x_k^t - x_k^{t+1}) U_{t+2} \end{aligned} \quad (\text{D.24})$$

$$= \frac{1}{N} \sum_{j \neq i} \sum_{p=2}^{aN} \sum_{\kappa \in A_p} \lambda_\kappa \xi_i^\kappa \xi_j^\kappa (x_j^{t+2,(p)} + \lambda_\kappa (x_j^t - x_j^{t+1}) U_{t+2}) \quad (\text{D.25})$$

$$\begin{aligned} &= \frac{1}{N} \sum_{j \neq i} \sum_{p=2}^{aN} \sum_{\kappa \in A_p} \lambda_\kappa \xi_i^\kappa \xi_j^\kappa (x_j^{t+2,(p)} \\ &\quad + \lambda_\kappa (x_j^t - (x_j^{t+1,(p)} + \lambda_\kappa U_{t+1} x_j^t)) U_{t+2}) \end{aligned} \quad (\text{D.26})$$

$$\begin{aligned} &= \frac{1}{N} \sum_{j \neq i} \sum_{p=2}^{aN} \sum_{\kappa \in A_p} \lambda_\kappa \xi_i^\kappa \xi_j^\kappa (x_j^{t+2,(p)} \\ &\quad - \lambda_\kappa U_{t+2} x_j^{t+1,(p)} + \lambda_\kappa U_{t+2} (1 - \lambda_\kappa U_{t+1}) x_j^t) \end{aligned} \quad (\text{D.27})$$

となる。よって、それぞれ

$$z_i^{t+1,(\kappa)} = \frac{1}{N} \sum_{j \neq i} \sum_{p=2}^{aN} \lambda_\kappa \xi_i^\kappa \xi_j^\kappa (x_j^{t+1,(p)} + \lambda_\kappa x_j^t U_{t+1}) \quad (\text{D.28})$$

$$z_i^{t+1} = \sum_{\kappa \in A_p} z_i^{t+1,(\kappa)} \quad (\text{D.29})$$

$$\begin{aligned} z_i^{t+2,(\kappa)} &= \frac{1}{N} \sum_{j \neq i} \sum_{p=2}^{aN} \lambda_\kappa \xi_i^\kappa \xi_j^\kappa (x_j^{t+2,(p)} \\ &\quad - \lambda_\kappa U_{t+2} x_j^{t+1,(p)} + \lambda_\kappa U_{t+2} (1 - \lambda_\kappa U_{t+1}) x_j^t) \end{aligned} \quad (\text{D.30})$$

$$z_i^{t+2} = \sum_{\kappa \in A_p} z_i^{t+2,(\kappa)} \quad (\text{D.31})$$

と書き直すことができる。これを使って、 $\rho_{t+1} = (E[(z_t - z_{t+1})^2])^{\frac{1}{2}}$, $\sigma_{t+2} = (E[z_{t+2}^2])^{\frac{1}{2}}$ を計算すればよい。そのためには、分散 $S_{t,t} = E[z_t, z_t]$, $S_{t+1,t+1} = E[z_{t+1}^2]$, $S_{t+2,t+2} = E[z_{t+1}^2]$ 、共分散 $S_{t,t+1} = E[z_t z_{t+1}]$ を求める必要がある。ここで、 $S_{t,t}^{(\kappa)} = E[(z_t^{(\kappa)})^2]$, $S_{t+1,t+1}^{(\kappa)} = E[(z_{t+1}^{(\kappa)})^2]$, $S_{t+2,t+2}^{(\kappa)} = E[(z_{t+1}^{(\kappa)})^2]$, $S_{t,t+1}^{(\kappa)} = E[z_t^{(\kappa)} z_{t+1}^{(\kappa)}]$ とすると、 $z^{(\kappa)}, z^{(\eta)}$ ($\kappa \neq \eta$) の間には相関が無いので、

$$S_{t,t} = \sum_{\kappa \in A} S_{t,t}^{(\kappa)} \quad (D.32)$$

$$S_{t,t+1} = \sum_{\kappa \in A} S_{t,t+1}^{(\kappa)} \quad (D.33)$$

$$S_{t+1,t+1} = \sum_{\kappa \in A} S_{t+1,t+1}^{(\kappa)} \quad (D.34)$$

$$S_{t+2,t+2} = \sum_{\kappa \in A} S_{t+2,t+2}^{(\kappa)} \quad (D.35)$$

が成り立つ。ここで $S_{t,\tau}^{(\kappa)}$ を具体的に計算すると、それぞれ

$$S_{t,t}^{(\kappa)} = E\left[\left(\frac{1}{N} \sum_{j \neq i} \sum_{p=2}^{aN} \lambda_{\kappa} \xi_i^{\kappa} \xi_j^{\kappa} x_j^{t+1,(p)}\right)^2\right] \quad (D.36)$$

$$= a\lambda_{\kappa}^2 Q_{t,t} \quad (D.37)$$

$$S_{t,t+1}^{(\kappa)} = E\left[\left(\frac{1}{N} \sum_{j \neq i} \sum_{p=2}^{aN} \lambda_{\kappa} \xi_i^{\kappa} \xi_j^{\kappa} x_j^t\right) \left(\frac{1}{N} \sum_{j \neq i} \sum_{p=2}^{aN} \lambda_{\kappa} \xi_i^{\kappa} \xi_j^{\kappa} (x_j^{t+1,(p)} + \lambda_{\kappa} U_{t+1} x_j^t)\right)\right] \quad (D.38)$$

$$= a\lambda_{\kappa}^2 (Q_{t,t+1} + \lambda_{\kappa} U_{t+1} Q_{t,t}) \quad (D.39)$$

$$= a\lambda_{\kappa}^2 Q_{t,t+1} + \lambda_{\kappa} U_{t+1} S_{t,t}^{(\kappa)} \quad (D.40)$$

$$S_{t+1,t+1}^{(\kappa)} = E\left[\left(\frac{1}{N} \sum_{j \neq i} \sum_{p=2}^{aN} \lambda_{\kappa} \xi_i^{\kappa} \xi_j^{\kappa} x_j^{t+1,(p)} + \lambda_{\kappa} U_{t+1} z_i^t\right)^2\right] \quad (D.41)$$

$$= E\left[\left(\frac{1}{N} \sum_{j \neq i} \sum_{p=2}^{aN} \lambda_{\kappa} \xi_i^{\kappa} \xi_j^{\kappa} (x_j^{t+1,(p)} + \lambda_{\kappa} U_{t+1} x_j^t)\right)^2\right] \quad (D.42)$$

$$= a\lambda_{\kappa}^2 (Q_{t+1,t+1} + 2\lambda_{\kappa} U_{t+1} Q_{t,t+1} + \lambda_{\kappa}^2 U_{t+1}^2 Q_{t,t}) \quad (D.43)$$

$$= a\lambda_{\kappa}^2 (Q_{t+1,t+1} + 2\lambda_{\kappa} U_{t+1} Q_{t,t+1}) + \lambda_{\kappa}^2 U_{t+1}^2 S_{t,t}^{(\kappa)} \quad (D.44)$$

$$S_{t+2,t+2}^{(\kappa)} = E\left[\left(\frac{1}{N} \sum_{j \neq i} \sum_{p=2}^{aN} \lambda_{\kappa} \xi_i^{\kappa} \xi_j^{\kappa} x_j^{t+2,(p)} + \lambda_{\kappa} U_{t+2} (z_i^t - z_i^{t+1})\right)^2\right] \quad (D.45)$$

$$= E\left[\left(\frac{1}{N} \sum_{j \neq i} \sum_{p=2}^{aN} \lambda_{\kappa} \xi_i^{\kappa} \xi_j^{\kappa} (x_j^{t+2,(p)} + \lambda_{\kappa} U_{t+2} x_j^t - \lambda_{\kappa} U_{t+2} (x_j^{t+1,(p)} + \lambda_{\kappa} U_{t+1} x_j^t))\right)^2\right] \quad (D.46)$$

$$= E\left[\left(\frac{1}{N} \sum_{j \neq i} \sum_{p=2}^{aN} \lambda_{\kappa} \xi_i^{\kappa} \xi_j^{\kappa} (x_j^{t+2,(p)} - \lambda_{\kappa} U_{t+2} x_j^{t+1,(p)}\right.\right.$$

$$+ \lambda_\kappa U_{t+2} (1 - \lambda_\kappa U_{t+1}) x_j^t \Big)^2 \Big] \quad (\text{D.47})$$

$$= a \lambda_\kappa^2 \left(Q_{t+2,t+2} + \lambda_\kappa^2 U_{t+2}^2 Q_{t+1,t+1} - 2 \lambda_\kappa U_{t+2} Q_{t+1,t+2} \right. \\ \left. + 2 \lambda_\kappa U_{t+2} (1 - \lambda_\kappa U_{t+1}) Q_{t,t+2} - 2 \lambda_\kappa^2 U_{t+2}^2 (1 - \lambda_\kappa U_{t+1}) Q_{t,t+1} \right. \\ \left. + \lambda_\kappa^2 U_{t+2}^2 (1 - \lambda_\kappa U_{t+1})^2 Q_{t,t} \right) \quad (\text{D.48})$$

$$= a \lambda_\kappa^2 \left(Q_{t+2,t+2} + \lambda_\kappa^2 U_{t+2}^2 Q_{t+1,t+1} - 2 \lambda_\kappa U_{t+2} Q_{t+1,t+2} \right. \\ \left. + 2 \lambda_\kappa U_{t+2} (1 - \lambda_\kappa U_{t+1}) Q_{t,t+2} - 2 \lambda_\kappa^2 U_{t+2}^2 (1 - \lambda_\kappa U_{t+1}) Q_{t,t+1} \right) \\ \left. + \lambda_\kappa^2 U_{t+2}^2 (1 - \lambda_\kappa U_{t+1})^2 S_{t,t}^{(\kappa)} \right) \quad (\text{D.49})$$

となる。ここで、 $Q_{t,\tau} = E[x_t x_\tau]$ である。この S を使って

$$\rho_{t+1}^2 = S_{t,t} + S_{t+1,t+1} - 2S_{t,t+1} \quad (\text{D.50})$$

$$\sigma_{t+2,t+2}^2 = S_{t+2,t+2} \quad (\text{D.51})$$

とノイズ項の分散を計算することができる。

しかし、これらを実際求めるには、以上の計算結果で現れた U と Q を得なくてはならない。まず、 U は

$$U_{t+1} = \frac{1}{\sigma_t} \int Dz z \langle \xi f_1(\xi m_t + \sigma_t z) \rangle_\xi \quad (\text{D.52})$$

$$U_{t+2} = \frac{1}{\sigma_s} \int Dz z \langle \xi f_0(\xi(m_t - m_{t+1}) + \rho_{t+1} z) \rangle_\xi \quad (\text{D.53})$$

と求められる。

次に $Q_{t,t}$, $Q_{t+1,t+1}$, $Q_{t+2,t+2}$, $Q_{t,t+1}$, $Q_{t,t+2}$, $Q_{t+1,t+2}$ を求める。まず、 $t \neq \tau$ の場合を考える。 x の間で相関を計算する時、前の時刻における z の相関を考慮する必要がある。しかし、 $Q_{t,t+1}$, $Q_{t,t+2}$ の場合は、 $t-1$ 以前の相関を時間は考えなくてよいので、 z の相関を考慮する必要はない。よって、問題になるのは $Q_{t+1,t+2}$ だけである。この計算を行うと

$$Q_{t+1,t+2} = \int dz_a \int dz_b \int dz_c \langle f_1 \left(\sum_{\mu \in A_1} \xi^\mu m_t^\mu + \sigma_t (\sigma_t^{(\alpha)} z_a + \sigma_t^{(\beta)} z_b) \right. \right. \\ \left. \left. + \Gamma_t \right) \times f_0 \left(\sum_{\mu \in A_1} \xi^\mu (m_t^\mu - m_{t+1}^\mu) \right. \right. \\ \left. \left. + \text{sgn}(C_{t,t+1}) \rho_{t+1} (\sigma_t^{(\alpha)} z_a + \sigma_t^{(\beta)} z_c) + \Gamma_{t+1} \right) \rangle_\xi \quad (\text{D.54})$$

$$C_{t,t+1}^{(\kappa)} = a \lambda_\kappa^2 (1 - \lambda_\kappa U_{t+1}) Q_{t,t+1}^{(\kappa)} + \lambda_\kappa U_{t+1} S_{t,t} \quad (\text{D.55})$$

$$C_{t,t+1} = \sum_{\kappa \in A} C_{t,t+1}^{(\kappa)} \quad (\text{D.56})$$

$$\sigma_t^{(\alpha)} = \sqrt{\frac{|C_{t,t+1}|}{\sigma_t \rho_{t+1}}} \quad (\text{D.57})$$

$$\sigma_t^{(\beta)} = \sqrt{1 - \frac{|C_{t,t+1}|}{\sigma_t \rho_{t+1}}} \quad (\text{D.58})$$

が導かれる。他の $Q_{t,\tau}$ は $Q_{t,\tau} = m_t^1 m_\tau^1$ と計算される。次に、 $t = \tau$ の場合だけ考える。 $Q_{t,t}$ は上の議論から $Q_{t,t} = \frac{1}{a} S_{t,t}$ で与えられ、 $Q_{t+1,t+1}, Q_{t+2,t+2}$ は

$$Q_{t+1,t+1} = \int dz \langle f_1(\sum_{\mu \in A_1} \xi^\mu m_t^\mu + \sigma_t z)^2 \rangle_\xi \quad (\text{D.59})$$

$$Q_{t+2,t+2} = \int dz \langle f_0(\sum_{\mu \in A_1} \xi^\mu (m_t^\mu - m_{t+1}^\mu) + \rho_{t+1} z)^2 \rangle_\xi \quad (\text{D.60})$$

のように求めることができる。この結果をまとめると統計神経力学のダイナミクスが得られる。

以上の議論では、ノイズ項の平均は0と仮定していたが、ノイズ項の平均も考慮した場合を考える。ノイズ項を平均まで考えると

$$z_i^{t+1} = \frac{1}{N} \sum_{j \neq i} \sum_{p=2}^{aN} \sum_{\kappa \in A_p} \xi_i^\kappa \xi_j^\kappa x_j^{t+1,(p)} + U_{t+1} z_i^t + a x_i^t U_{t+1} \quad (\text{D.61})$$

$$z_i^{t+2} = \frac{1}{N} \sum_{j \neq i} \sum_{p=2}^{aN} \sum_{\kappa \in A_p} \xi_i^\kappa \xi_j^\kappa x_j^{t+2,(p)} + U_{t+2} (z_i^t - z_i^{t+1}) + a (x_i^t - x_i^{t+1}) U_{t+2} \quad (\text{D.62})$$

となる。本来は i によって x は異なるが、これを平均化し、全ての i について一様に m^1 で近似すれば、上の式は

$$z_i^{t+1} = \frac{1}{N} \sum_{j \neq i} \sum_{p=2}^{aN} \sum_{\kappa \in A_p} \xi_i^\kappa \xi_j^\kappa x_j^{t+1,(p)} + U_{t+1} z_i^t + a m^1 U_{t+1} \quad (\text{D.63})$$

$$z_i^{t+2} = \frac{1}{N} \sum_{j \neq i} \sum_{p=2}^{aN} \sum_{\kappa \in A_p} \xi_i^\kappa \xi_j^\kappa x_j^{t+2,(p)} + U_{t+2} (z_i^t - z_i^{t+1}) + a (m_t^1 - m_{t+1}^1) U_{t+2} \quad (\text{D.64})$$

となる。このノイズ項の平均

$$\Gamma_{t+1} = a m_t U_{t+1} + U_{t+1} \Gamma_t \quad (\text{D.65})$$

$$\Gamma_{t+2} = a (m_t - m_{t+1}) U_{t+2} + U_{t+2} (\Gamma_t - \Gamma_{t+1}) \quad (\text{D.66})$$

を上記の計算結果に導入すると、本文中の式 (6.51)～(6.72) を得ることができる。

発表文献

- <1> 掛谷英紀 (1995): 連想記憶システムの研究 – 記憶行列の固有空間解析とその応用 –, 東京大学大学院工学系研究科計数工学専攻修士論文.

学術論文(既刊)

- <2> 掛谷英紀, 金道敏樹 (1994): 符号反転記憶法による自己相関連想記憶の改良, 日本神経回路学会誌, **1-1**, 20-26.
- <3> Takeya, H. and Kindo, T. (1996): Hierarchical concept formation in associative memory composed of neuro-window elements, *Neural Networks*, **9**, 1095-1098.
- <4> Takeya, H. and Kindo, T. (1997): Eigenspace separation of autocorrelation memory matrices for capacity expansion, *Neural Networks*, **10**, 833-843.

学術論文(印刷中)

- <5> Kindo, T. and Takeya, H.: A geometrical analysis of associative memory, *Neural Networks*, in press.

学術論文(投稿中)

- <6> Takeya, H., Kindo, T., and Okabe, Y.: Geometry of neuro-window associative memory, *submitted to Neural Networks*.
- <7> Takeya, H. and Okabe, Y.: Selective retrieval of memory and concept sequences through neuro-windows, *submitted to IEEE Trans. NN*.

学会予稿、技術報告

- <8> 掛谷英紀, 金道敏樹 (1994): 自己相関行列による連想記憶における記憶法の改良, 信学技報, **NC94-25**, 9-16.
- <9> Takeya, H. and Kindo, T. (1994): Analysis and improvement of associative memory from the viewpoint of linear algebra, *Intelligent Engineering Systems Through Artificial Neural Networks*, **4**, 121-126.

- <10> 金道敏樹, 掛谷英紀, 森沢茂 (1994): 連想記憶の想起過程の幾何的解析, 電子情報通信学会秋期大会プロシーディング, 6, 451-452.
- <11> 掛谷英紀, 金道敏樹, 中野馨 (1994): 時系列情報処理における部分反転ダイナミクスの応用, 日本神経回路学会第5回全国大会講演論文集, 200-201.
- <12> 掛谷英紀, 金道敏樹 (1995): ニューロウインドウ法を用いた連想記憶, 信学技報, NC94-117, 25-32.
- <13> 金道敏樹, 掛谷英紀 (1995): 自己相関連想記憶の幾何的な性質, 信学技報, NC94-116, 17-24.
- <14> 掛谷英紀, 金道敏樹, 中野馨 (1995): 階層的制御機構を持つ神経回路網モデル, RWC 情報統合ワークショップ '95 プロシーディング, 41-48.
- <15> Kakeya, H., Kindo, T., and Nakano, K. (1995): Neural network models with hierarchical control mechanisms, *Proc. of Information Integration Workshop (TR-95010)*, 32-40.
- <16> 掛谷英紀, 金道敏樹, 岡部洋一 (1995): 非単調ニューロンを用いた相関学習が構成する記憶空間の性質, 日本神経回路学会第6回全国大会講演論文集, 103-104.
- <17> 掛谷英紀 (1995): 連想記憶から見た記号とパターン, 電子情報通信学会「若手セミナー: 記号とパターンの統合」予稿集, 78.
- <18> 金道敏樹, 岡夏樹, 掛谷英紀 (1995): 連想記憶の幾何的理論による心理現象の説明の試み - ロッシュの基本レベルを題材にして -, 認知科学会「学習と対話」, 95-2, 24-33.
- <19> 掛谷英紀, 金道敏樹, 岡田真人, 岡部洋一 (1996): 連続時間型ニューロウインドウ法の解析, 信学技報, NC95-131, 121-128.
- <20> 金道敏樹, 掛谷英紀 (1996): 忘却を行なう自己相関連想記憶の幾何的性質, 信学技報, NC95-133, 137-144.
- <21> 掛谷英紀, 金道敏樹, 岡部洋一 (1996): 系列の連想におけるニューロウインドウ法, 日本神経回路学会第7回全国大会講演論文集, 255-256.
- <22> 掛谷英紀, 岡部洋一 (1996): 論理関数をBP学習した多層パーセプトロンの中間層発火率, 日本神経回路学会第7回全国大会講演論文集, 98-99.
- <23> 掛谷英紀, 岡田真人, 岡部洋一 (1996): ニューロウインドウ連想記憶のS/N解析, 日本神経回路学会第7回全国大会講演論文集, 49-51.

- <24> 掛谷英紀, 岡部洋一 (1997): 離散同期型ダイナミクスを用いた組み合わせ最適解の探索, 信学技法, **NPL96-125/NC96-79**, 9-16.
- <25> Kakeya, H. and Okabe, Y. (1997): Capacity expansion and concept formation in sequential associative memory, *Proc. of ICNN '97*, **2**, 1332-1337.
- <26> Kakeya, H. and Okabe, Y. (1997): Parallel and synchronous search for combinatory quasi-optimum solutions, *Proc. of ICNN '97*, **2**, 1284-1288.
- <27> 掛谷英紀, 岡田真人, 岡部洋一 (1997): 部分反転法の統計神経力学, 信学技報, **NC96-127**, 111-118.
- <28> 掛谷英紀, 岡部洋一 (1997): 並列同期デジタル計算における組み合わせ最適解の高速探索, 電気学会情報処理研究会資料, **IP97**, 1-6.
- <29> 井上浩一, 掛谷英紀, 岡部洋一 (1997): (0,1)出力型ニューロウインドウ連想記憶, 日本神経回路学会第8回全国大会講演論文集, 238-239.
- <30> 掛谷英紀, 岡田真人, 岡部洋一 (1997): ニューロウインドウ連想記憶の統計神経力学, 日本神経回路学会第8回全国大会講演論文集, 232-233.
- <31> 掛谷英紀, 岡部洋一 (1997): 並列同期計算による巡回セールスマン問題の高速解法, 日本神経回路学会第8回全国大会講演論文集, 31-32.
- <32> 掛谷英紀, 岡部洋一 (1997): 非対称荷重を持つニューラルネットワークの幾何的解析, 信学技報, **NC97-57**, 17-24.