



1993年  
267

学位論文

神経系の自己組織的形成に関するモデルの  
数理的解析

1992年12月21日 提出

指導教官 齋藤正男 教授

東京大学大学院 工学系研究科

電子工学専攻

07091

田中 利幸

# 目次

1 序論	1
1.1 神経系の研究と神経回路網理論	2
1.2 神経系の自己組織的形成	5
1.3 本論文の目的	7
1.4 本論文の構成	10
2 生体の神経系に関する基礎的事項	11
2.1 本章の目的	12
2.2 ニューロン	12
2.3 神経系の学習と自己組織化の機構	14
2.3.1 シナプスの可塑性	14
2.3.2 Hebb 則	15
2.4 初期視覚系	16
2.4.1 概要	16
2.4.2 網膜	16
2.4.3 外側膝状体	18
2.4.4 一次視覚野	18
2.5 考察	20
図表	23

3	神経回路網理論に関する基礎的事項	29
3.1	本章の目的	30
3.2	形式ニューロン	30
3.3	Hebb 則	33
3.4	神経場	34
3.5	考察	35
	図表	36
4	構造の大域的な特徴の自己組織的形成に関するモデル	39
4.1	本章の目的	40
4.2	Kohonen のモデル	41
4.2.1	基本的な競合学習	41
4.2.2	協調学習	42
4.2.3	神経場モデルおよび神経系との関連	43
4.3	平均学習方程式	46
4.4	参照ベクトルの漸近的分布	48
4.4.1	モデルが 1 次元の場合	48
4.4.2	モデルが $n$ 次元の場合	49
4.4.3	計算機シミュレーション	53
4.5	学習の収束性	54
4.5.1	概要	54
4.5.2	学習方程式の流体近似	55
4.5.3	Lyapunov 関数	56
4.6	参照ベクトルの整列に関する検討	57
4.7	考察	59
	図表	64

5	構造の局所的な特徴の自己組織的形成に関するモデル	73
5.1	本章の目的	74
5.2	Linsker のモデル	74
5.2.1	基本的な学習の機構	74
5.2.2	平均学習方程式	75
5.3	on-中心形 / off-中心形ニューロンの形成	76
5.4	方位選択性ニューロンの形成	77
5.5	考察	78
	図表	81
6	神経系の自己組織的形成に関する考察	88
7	結論	92
	謝辞	94
	参考文献	95
	A 式 (4.28) の導出	106
	B 式 (4.38) の導出	109
	C 定理の証明	111

## 第 1 章

### 序論

形式主義は、諸学問・芸術において異なった意味をもっており、またときには異なる名称でよばれている。たとえば数学におけるヒルベルトの形式主義と文学批評における形式主義は、<sup>フォルマリスム</sup>意味がちがっているし無関係なものとみなされている。さらに、形式主義は、ある領域では<sup>シンボリック</sup>記号論理学とよばれ、またべつの諸領域では、<sup>セミオティックス</sup>記号論、あるいはサイバネティックスとよばれたりしている。 ([44])

## 1.1 神経系の研究と神経回路網理論

サイバネティクスとは、いうまでもなく工学における形式主義につけられた名前である。形式主義は、多くの学問領域において複雑でそれまで手がつけられなかった問題を手際よく取り扱う道を開いた。サイバネティクスによって工学に開かれた新しい分野の一つが生体工学である。言いかえれば、生体工学というのは、生体を形式的なひとつの工学システムとして捉えることでそれを理解しようとする立場のことである。

人間をひとつの工学システムとして見てみると、それは多くの器官が複雑、巧妙に組み合わさり、統合された精緻なシステムである。人間の各器官の基本的なはたらきとつくり、およびそれら機能と構造との関係については、今日ではサイバネティクスによらなくても概ねよく理解されているとあってよい。だがこのことは、脳をはじめとする生体の神経系に関しては当てはまらないように思われる。脳について現在までにわかっていることがらを集めてみても、脳の基本的なはたらきとつくりを簡潔に言いあらわすことはできないし、将来にわたっても絶望的だろう。脳のこの複雑さのために、脳の研究では形式主義的方法が積極的に使われている。

脳を初めて明確に工学システムとして扱ったのは、Wiener のサイバネティクスにおいてであろう。サイバネティクスの初期において、脳は「論理機械」という形式で扱われている [95]。この取扱いは、McCulloch と Pitts の成果 — いかなる有限の論理も、McCulloch-Pitts の形式ニューロンによって実現できる [68] — に強く影響をうけている。それ以来、サイバネティクスにおいては、脳はまず情報処理システムとして捉えられるのが普通である。工学システムとしての脳の取扱いに対して、McCulloch-Pitts の成果が与えた影響はそれだけではない。明らかに実際のニューロンのはたらきは McCulloch-Pitts の形式ニューロンよりもはるかに複雑だが、彼らの仕事は、その複雑なニューロンを形式化し、それら形式ニューロンが組み合わさってできる「神経回路網」の水準で神経系の機能を捉えようとする、という態度を生み出し、定着させた。今日「神経回路網理論」とよばれる領域が形成されたのは、ここにおいてであるといつてよい。言いかえれば、神経回路網理論とは、工学の立場からの神経系の形式主義的な研究で、とくにニューロンの水準で形式化をおこない回路網の水準での情報処理システムとしての神経系のはたらきを理解しようとする立場のことである。

生体の神経系のはたらきを理解するためのアプローチは、何も神経回路網理論だけに限られるわけではない。例えば、脳の構造は階層的だと考えられているから、脳を考えるときには、階層のそれぞれの水準で、その水準に適した問題を設定することができる。脳の構造の階層としては、例えば分子の水準、シナプスの水準、ニューロンの水準、回路網の水準、層構造の水準、機能地図の水準、システムの水準、というようなものが考えられるだろう。そのなかで神経回路網理論は、いうまでもなく回路網の水準からのアプローチに相当する。また、分子生物学や神経生理学はそれより下の水準からのアプローチということになるし、認知心理学はより上の水準に注目したアプローチということになる。

ここでみたような、脳の研究にはいろいろな水準からのアプローチが可能であり、また必要である、という考え方は、David Marr [65] によってはじめて研究の方法論として明確に意識されたといわれている。Marr は、視覚情報処理の問題を計算理論の立場から扱うという作業を通して、この考え方を具体的に展開してみせた。彼自身は研究の水準として計算理論、アルゴリズム、実装、の三つを設定しているが、彼の研究に触発された、のちの研究(例えば、川人ら [45,46] など)をみると、Marr の視覚研究はそれ全体として回路網の水準より上位の水準での脳研究として位置づけられることがはっきりする。

Marr が視覚研究で採用した方法論は、回路網の水準より上位の、特定の水準に限定されるような特殊なものでは決してない。むしろもっと一般化して、脳の階層的構造のどの水準でも、計算理論、アルゴリズム、実装のそれぞれに関する問題が存在するというべきである。例えば、単一のニューロンが行なう計算は極めて複雑なものでありうるし、人間が習得できる言語を猿は習得できないという事実からは、システムの水準での言語機能の実装という問題が提起されうるだろう。

Marr はまた、研究の各水準は比較的独立に扱われるべきで、ある特定の問題はそれに適した水準で説明されるべきである、と主張する。この主張は、脳の研究を階層的にみる考え方をさらに押し進めたものといえる。Marr 自身の視覚情報処理の研究をみると、「我々は視覚的に物体の 3 次元形状を認識できる」という事実が強調され、そこからのトップダウン的方法が主に採られているように見える。けれども、例えばニューロ

ンの水準の知見は無視されているかということそうではない。Churchland と Sejnowski はむしろ、この種のボトムアップ的な知識が、研究の方向性を決める上で効果的に利用されていることを指摘している [10]。

したがって、神経系の研究に関していえば、取り扱う問題に対応した適切な水準を選び、その上下の水準に関する知見をどちらも適切に取り込んで形式化する、という、いわばトップダウン/ボトムアップ併用形の方法論が重要であるということになる。ただし、ある特定の形式化がなされたときに、それが「適切」なものであるかどうか、ということについては、現段階では一般的に判定する方法があるわけではないから、それは個別に判断されなければならない [10]。いいかえれば、神経系の個々の研究を統一的な立場からまとめあげることは今のところできないので、神経系の研究全体のなかで個々の研究を位置づけることは、トップダウン/ボトムアップ的な各水準の間のやりとりのなかでなされていかなければならない、ということである。

神経系を情報処理システムとして捉えるなら、工学あるいはサイバネティクスの方法論は神経系のしくみを説明するのに適している。けれどもそれだけではなく、工学は神経系との関連に関わりなく情報処理システムを研究対象としている。そこで、神経回路網のような生体の神経系のモデルを、生体の神経系からはなれて、人工的な情報処理システムのひとつの形態として研究しようとする立場もある。このような研究の分野は、neurocomputing といわれて生体の神経系を工学的な情報処理システムとして研究する分野 computational neuroscience と便宜的に区別されることがある [4]。

今日、工学の分野での情報処理システムとしては、ノイマン形コンピュータが高度な発展を遂げている。このノイマン形コンピュータとの対比によって、今日の神経回路網理論では、脳という情報処理システムは以下のような特徴をもつものとして位置づけられている。

- (a) 並列分散形の情報処理
- (b) 冗長なシステム構成による頑健性
- (c) 学習, 適応, 自己組織化の能力
- (d) あいまいな情報処理



これらの特徴はノイマン形コンピュータが苦手とするものであり、人工的な情報処理システムでこれらを実現することは、情報処理の分野でも重要な課題である。脳など神経系のもつこれらの特徴を工学的に理解することは、人工的な情報処理システムへその知識を応用することをも可能にする。実際に、神経回路網理論の研究成果のいくつかは、適当な情報処理課題へ応用されて成果をあげている。

このように、computational neuroscience の立場と neurocomputing の立場とは便宜的にわけて考えることもできるけれども、前者における研究成果は直ちに後者において適用しうるし、また逆に後者の結果が前者の研究に示唆を与えることもある、というように、相互に影響しあい、自由に入れ替わるような性格のものである。理論が形式化されていることが、このような往来を可能にしている。

## 1.2 神経系の自己組織的形成

本論文で取り扱うのは、神経回路網理論の水準での、神経系の自己組織的形成に関する問題である。

生体の神経系の構造に関しては、神経生理学的な研究が精力的に行われている。なかでも感覚神経系、とくに視覚系がよく研究されているようである。視覚系を例にとると、そこにはいくつもの特徴的な構造がみつまっている。さらにそれらのうちのいくつかは、視覚情報処理に適した構造であるという意味づけがなされている。このように、普遍的にみられる構造であり、さらに機能という面からみても意味づけがなされているような構造については、「神経系のこのようないま構造はどうやって形成されるのか」という問題が、神経回路網理論の立場から提起される。

「神経系の構成は細かいところに至るまで、すべて遺伝で決められている」というのはひとつの考え方である。けれども、人間の脳を例にとると、ニューロンの数がおよそ  $10^{12}$ 、シナプスの数はおよそ  $10^{15}$  であると見積もられている [10]。これだけの情報を遺伝ですべてまかなう、というのは不自然であるし、またあらかじめ決められているとするのは、ある意味で柔軟性に欠けるように思われる。事実、生後のある特定の期間に異常な視覚環境におかれると、通常みられるような視覚系の構造が形成されなくなる、といったことが知られているので、遺伝のほかには何らかの形成機構があることは疑いが

ない。これらのことから、神経系の構成が形づくられる際には、環境からの刺激に対応して神経系が自らをつくりあげる、いわゆる自己組織化の機構が重要である、という考え方が現れ、現在では広く受け入れられている。この考え方では、遺伝のはたらきとしては、自己組織化の機構を定め、神経系の巨視的な構成を条件づける、というような大まかな役割が想定される。

神経系の自己組織的形成に関する、工学の立場からの研究について考える。神経系が自己組織的に形成される際に、「何が必要ならなければならないか」という、計算理論の水準での問題の記述を考えてみると、これはそれほど複雑なものではない。神経系のある構造や機能が実際に存在することが神経解剖学的にあるいは神経生理学的にわかってきたとき、そのような構造や機能が、完全に遺伝によるのではなく自己組織的に形成されなければならない、というのが、計算理論の水準での問題の記述である。

また、神経系の構造や機能が自己組織的に形成されるものであるなら、通常と異なる環境のもとで形成される構造や機能は、通常環境のもとで形成される構造や環境とは異なってくるはずである。そのような実験事実が得られたとき、そのような異常な環境では、同じ自己組織的形成の原理にもとづく神経系は実験事実に合うような構造を形成しなければならない。このことも、神経系の自己組織的形成に関する問題の計算理論の水準における記述の一部をなしている。

次に、実装の水準での、神経系の自己組織的形成に関する問題の記述を考える。すでに述べたように、神経回路網理論の立場では、おもにニューロンの水準で形式化を行ない、神経系を形式ニューロンからなる回路網として取り扱う。通常使われる形式ニューロンは、多入力一出力の素子である。入力の線形加重和をとり、ある出力関数にその和を通した結果を出力する、というような、単純な動作をする。神経系の自己組織的形成の研究では、神経系が環境に応じてその性質を変化させていく、すなわちある意味で学習を行なう様子を取り扱うのだが、この神経系の変化というものは、個々の形式ニューロンが、一つ一つの入力の重みを調節するという形で自身の性質を変化させる、という形で取り扱われる。

このような取り扱いのなかでよく知られているものとして、Hebb 則 [28] がある。Hebb 則は、二つのニューロン間の結合が、その二つのニューロンの活動をそろえる

ように変化する、というごく単純な形の規則である。もちろん、神経系が二つのニューロンの間の結合を変化させるやりかたは現実には極めて複雑であるといわれている。けれども、Hebb 則に近い方式が実際にもよく観察されていて、基本的な考え方としては Hebb 則は妥当なものとして支持されている。

神経系の自己組織的形成を実装の水準からみれば、以上で述べたように単純な動作をする素子が Hebb 則のような単純な学習規則に従うとき、実際に観察されるような神経系の構造に対応するようなものができなければならない。これが、実装の水準での、神経系の自己組織的形成に関する問題の記述である。

最後に、計算理論の水準と実装の水準とをつなぐ、アルゴリズムの水準について述べる。実際に観察されるような神経系の構造については、それをつくりあげる仕組みがなければならぬ。一方で、神経回路網理論の立場からは、神経系は、単純な動作をするニューロンからなる回路網とみなされるのだが、ニューロンの単純な動作から、実際に観察されるような構造の形成を説明できなければならない。この両者のあいだをつなぐ説明を与えるのが、アルゴリズムの水準である。神経系の自己組織的形成に関する問題を、神経回路網の水準で取り扱う場合には、必然的にアルゴリズムの水準が主要なものになる。

### 1.3 本論文の目的

神経生理学、神経解剖学での研究によって、神経系の構造や機能に関するたくさんの知見が得られてきた。それらに呼応して、神経回路網理論の立場からもいくつものモデルが提案されてきた。これらのモデルの目的は、ひとつには複雑な実際の神経系が示す特徴的なふるまいを、その神経系がもつ本質的な性質は保ったままで単純化したモデルを通して理解することである。あるモデルが、神経系の示す特徴的なふるまいを再現できたときには、そのモデルはいま問題としているふるまいに関する、アルゴリズムの水準での記述を与えていると考えることができる。また、神経系が示す特徴的なふるまいが情報処理の立場からみて有用なものである場合には、そのアルゴリズムを表現している、本質を捉えた単純なモデルを活用することによって、人工的な情報処理システムへ

のアルゴリズムの実装を効率的に行なうことができる。これが、モデルのもうひとつの目的である。

これらのモデルはふつう、単純な動作規則に従う形式ニューロンから構成される回路網、という形をとる。そのおかげで、これらのモデルの動作規則を手続き的に理解するのはそれほど難しくはない。計算機上にモデルを実現し、実際に動作させてみることも容易である。モデルの動作を計算機シミュレーションによって確かめることが、しばしば行われている。実際、神経回路網理論の分野では、計算機シミュレーションはモデルの性質を理解するうえで重要な手段である。

その反面で、モデルの動作を理論的に理解するのは、かえって難しいのが普通である。理論的な取扱いが単純であるということと、手続きが単純であることとは違うからである。だから、あるモデルに関するある性質が経験的には広く受け入れられているが、理論的には必ずしも保証されていない、ということは、よくあることである。けれども、モデルの性質を理解するためには、計算機シミュレーションによるいくつかの結果だけでなく、理論的な基礎づけもまた不可欠である。

既に述べたように、感覚神経系、とくに視覚系に関しては神経生理学、神経解剖学の立場からの研究が進んでいて、それらの神経系での構造的特徴というべきものがわかってきている。視覚系の初期段階すなわち初期視覚系の構造が自己組織的に形成されることに関して、いくつかの神経回路網の立場からのモデルが提案されている。そのなかで本論文では、Kohonen が提案したモデル [49] および Linsker が提案したモデル [58-60] を取り扱う。

初期視覚系のもつ構造の特徴を、大域的な特徴と、局所的な特徴との二種類に分けて考えると、Kohonen のモデルは構造の大域的な特徴の形成に関するモデルであり、Linsker のモデルは構造の局所的な特徴の形成に関するモデルである、ということがいえる。

Kohonen のモデルが関連する、視覚系の構造の大域的な特徴とは、例えば以下のようなものである。初期視覚系は網膜にはじまる層構造を成しているとみなすことができるが、層間の結合は大域的にみればしばしば各層での位相を保存するような形で構成されている。例えば、網膜上の各点は大まかにいえば一次視覚野上の各点と連続的に対応

づけることができる。このような位相保存写像は、網膜部位再現あるいは retinotopy と呼ばれる。同様の構造は、他の感覚神経系、例えば体性感覚野や聴覚野でも存在するとされており、それぞれ somatotopy, tonotopy とよばれている。

さらに、これらの位相保存写像では、各点での写像の拡大率が、そこでの刺激の頻度に対応するような形になっている。例えば、網膜の中心窩の周辺は一次視覚野上で大きく拡大されているのに対し、周辺視に対応する部分はごく小さい面積を占めるに過ぎない。また、体性感覚野でも、指先に対応する部分は背中に対応する部分よりも遥かに大きい面積を占めている。このような性質を、各点での写像の拡大率が対応する感覚受容器での刺激の頻度にほぼ比例する、という経験則に基づいて、本論文では「(写像の拡大率に関する) 比例則」と呼んで参照することにする(感覚野の単位面積に対応する刺激の受容野の面積がそこでの刺激の頻度に逆比例する、というみかたをして、「逆比例則 (inverse proportionality)」と呼ばれることもある) Kohonen のモデルも、同様の性質を示すことが、計算機シミュレーションによって示されている。

いっぽう、初期視覚系の構造の局所的な特徴に注目すると、ひとつひとつのニューロンは網膜からの大域的写像できまるような、視覚刺激の位置に対する選択性を示すだけでなく、線分の方位や運動方向といった、視覚刺激の特定の特徴に対して選択性を示すことが観察される。このようなニューロンは、特徴検出細胞といわれる。Linsker のモデルは、このような特徴検出細胞の形成に関連するモデルであり、計算機シミュレーションによって、線分の方位を検出するようなニューロンが形成されることが示されている。

しかしながら、これらのモデルに関しても、計算機シミュレーションが先行していて、モデルの理論的な検討が遅れている。そこで本論文では、これらのモデルの性質を理論的に検討し、それを通して初期視覚系の構造の自己組織的形成に関する考察を行なうことを目的とする。

また、いうまでもなく本論文は工学の立場からの神経系の研究に属するものだが、工学の立場からの神経系の研究には、そこから得られる成果を人工的なシステムに応用しよう、という目的もある。神経系の自己組織的形成に関する研究でいえば、人工的な情報処理システムに対して適応能力や自律性をもたせる、ということが考えられる。実際

に Kohonen のモデルは、適応信号処理のシステムとしてすでに様々な分野で応用が進んでいる。このようなモデルを実際の問題に応用する際にも、モデルの理論的性質を理解することが重要である。本論文では、このような応用の側面に関連した考察も行なう。

## 1.4 本論文の構成

以下に本論文の構成を示す。第 2 章、第 3 章では、本論文での議論に関して基礎となる事項を述べる。第 2 章では、生体の神経系に関する知見を述べる。第 3 章では、神経回路網理論に関する基礎的な事項についてまとめる。形式ニューロン、Hebb 則、神経場の概念などを取り扱う。

第 4 章、第 5 章では、単純な Hebb 則にもとづく神経回路網モデルによって、いくつかの特徴的な構造が自己組織的に形成されることについて、理論的に解析をおこなう。

第 4 章では、感覚系の構造の大域的な特徴の自己組織的形成に関連する、Kohonen のモデルを取り扱う。はじめにモデルの説明を与え、実際の生体の神経系とのモデルの対応づけを議論する。モデルは、入力に対して適応的に構造を変化させるようなシステムとして定式化される。次に、モデルが形成する構造が、モデルに与えられた入力の性質とどのように関係するのかを理論的に解析する。また、モデルの収束性に関する議論もおこなう。

第 5 章では、視覚系の構造の局所的な特徴の自己組織的形成に関連する、Linsker のモデルを取り扱う。モデルの説明と、モデルの性質について現在までに知られている理論的な成果とを述べ、方位選択性ニューロンの形成に関して理論的な解析をおこなう。

第 6 章では、さらに高次の特徴に対して選択性を示すようなニューロンの形成が、Hebb 則によってどのように理解できるかを、視野運動選択性ニューロンを例にとって考察する。

第 7 章で、本論文の結論を述べる。

## 第 2 章

# 生体の神経系に関する基礎的事項

要するに、人間の解剖は猿の解剖にたいするひとつの鍵である。これに反して、低級な種類の動物にある、より高級な動物への暗示が理解されるのは、この高級なものそのものがすでに知られているばかりである。 ([66])

## 2.1 本章の目的

本章では、本論文の基礎となる、生体の神経系に関することがらについてまとめる。神経回路網理論の立場から神経系の自己組織的形成を取り扱う、というのが本論文の立場である。したがって、生体の神経系の構成要素としてはニューロンを考え、たくさんニューロンが結合することによって神経系が形成されている、というかたちで神経系をとらえる。また、神経系の各モジュールで実現されている機能が異なっているのは、構成要素としてのニューロンがモジュールごとに異なっているのではなく、ニューロンの結合様式が異なっているからだと考える。

このような前提で神経系とくに初期視覚系の自己組織的形成を取り扱うために、神経系の構成要素としてのニューロンと初期視覚系におけるニューロンの結合様式に関する知見、およびニューロンの結合様式が変化するときの機構すなわちシナプス可塑性について述べる。

## 2.2 ニューロン

ニューロン(神経細胞ともよばれる)は、筋線維などとともに、興奮性の細胞膜をもった細胞である。ニューロンの興奮は電気的なものであり、ニューロンが興奮すると、細胞膜内外の電位差が能動的に変化する。この電位変化は Hodgkin-Huxley の方程式に従う、決まった形での時間的変動を示す。ニューロンの興奮に伴うこのような電位変化は、活動電位といわれている。

細胞膜のどこかで活動電位が発生すると、その活動電位は細胞膜に沿ってニューロンの他の部分へも空間的に伝搬していく。ニューロンは、シナプスとよばれる構造を介して他のニューロンと結合しているが、活動電位が伝搬して行って、このシナプスに到達すると、シナプスからの化学物質の放出を引き起こす。この化学物質は、相手側のニューロンの細胞膜に対して、電位の変化を引き起こす性質をもっており、これによってひとつのニューロンの電氣的興奮が他のニューロンにも伝えられるので、神経伝達物質とよばれている。神経伝達物質にはいくつかの種類があることがわかっている。そのうちのあるものは、相手側のニューロンの細胞膜に対して正の向きの電位変化を引き起



こす。またあるものは、相手側のニューロンの細胞膜に対して負の向きの電位変化を引き起こす。前者は興奮性神経伝達物質とよばれ、後者は抑制性神経伝達物質とよばれる。なお、シナプスでのニューロンどうしの結合は、一方向性である(下等な動物では例外がある)。

ひとつのニューロンは、たくさんのニューロンからシナプスを介して細胞膜の電位変化という影響をうけとる。一般に、シナプスを通して引き起こされる細胞膜の電位変化は、ひとつひとつを取り出してみれば小さいものである。けれどもそれらは細胞膜上で加算され、やはり Hodgkin-Huxley 方程式であらわされるようなあるしきい値を電位が超えると、そこでまた活動電位が発生する。

このように、ニューロンは互いに電氣的に影響を与えたり受けたりしているので、神経系では活動電位が情報を担っていると考えられている。上で述べたように、活動電位の波形はそれがいったん発生してしまえば、どこでどのように発生したかに関係なく、いつも同じ波形を示す。したがって、ニューロンが表現する情報は、それが興奮したか興奮しないか、という、いわば 1 ビットの情報であるといわれている。活動電位のこのような性質は、悉無律とよばれている。

活動電位の悉無律に基づいて、McCulloch と Pitts とは、ニューロンを 0 か 1 かの 2 値を表現するものとして考えた [68] が、実際の神経系では、ひとつひとつのニューロンの動作は互いに同期がとれているわけではないから、実際に神経系で取り扱われている情報は活動電位のある、なしのような細かい情報ではないのではないか、と考えるのは自然である。ニューロンは、個々の活動電位ではなくて活動電位の時間的頻度という形で情報を表現している、という考え方が現在では一般的に受け入れられている。

実際には、例えばフクロウの音源定位 [19] などのように、ひとつひとつの活動電位のタイミングが情報を処理するうえで重要であることがわかっているケースもいくつか知られている。また、視覚系での情報処理でも、活動電位のタイミングが利用されるとする見解もある [14, 48, 82]。本論文では、神経系の自己組織的形成という、いわば時間的なスケールが非常に長い現象を取り扱うので、このような短い時間スケールでの現象は取り扱わず、ニューロンは情報を活動電位の時間的頻度という形で表現している、という考え方に従う。

また、ニューロンの情報表現が活動電位の時間的頻度という形であると考えたとしても、ひとつのニューロンが受け取った情報をどう処理して出力するかと考えるのか、という問題は検討する必要がある。この問題については古くから多くの研究がある [32]。とくに、短い時間スケールでの過渡的なふるまいに関しては、込み入った議論が必要になると考えられる [88, 89]。ここでもやはり、本論文で取り扱う現象が時間的なスケールの長いものであることから、この問題についてはこれ以上深く立ち入らない。長い時間スケールでは、ニューロンは受け取った活動電位の時間的頻度をシナプスの伝達効率で重みづけして加算し、その総和の関数として自分自身の活動電位の時間的頻度をきめる、というように考えられるのがふつうである。

## 2.3 神経系の学習と自己組織化の機構

### 2.3.1 シナプスの可塑性

生体の神経系では、ニューロンの数は生まれてからはほとんど変化しないことが知られている。ニューロンの数は、神経系の発生のある段階で急に増えるが、その後は細胞分裂をしなくなるために、ほとんど増加しない。むしろ加齢にともなって減少していく。いっぽう、神経系の機能は、ニューロンの数の増加が止まってからも、様々な変化を示す。感覚系の精緻な構造は出生前後に形成される。また、条件づけや学習とよばれるものは個体が十分成熟してからでも可能だが、これも神経系の機能の変化のひとつである。

神経系の構成要素であるニューロンの数は増加しないのに、神経系の機能は変化する。これは、ニューロンどうしの結合様式の変化によるものであると考えることができる。すなわち、神経系の機能の変化は、シナプスでの信号の伝達効率の変化に由来する、ということである。もちろん、神経系のはたらきはほとんどわかっていないといってよく、神経系の機能の変化がシナプスの伝達効率の変化によってどの程度説明されるものなのか、あるいは他に重要な機構があるのかどうか、というような点に関しては、まだよくわかっていない。けれども、シナプスの伝達効率の変化が、神経系の機能の変化に関する基本的な機構である、という考えは、現在では広く受け入れられていて、この考えを作業仮説として多くの研究が行われている。シナプスの伝達効率が可変であることは、シナプスの可塑性とよばれている。

### 2.3.2 Hebb 則

シナプスの可塑性に関して生理学的な学習規則という形ではじめて明確な主張をしたのは、Hebb であるとされている。シナプスの可塑性に関して Hebb が行った主張は、以下のようなものである [28].

さて、反響性の活動(あるいは「痕跡」)が持続したり繰り返されたりすると、細胞においてそれを安定にするような永続的な変化が引き起こされる、と仮定しよう。この仮定は、以下のように精密に表現することができる: 細胞 A の軸索が細胞 B を興奮させられるくらい近くにあつて、細胞 B の活動電位の発生に繰り返し執拗に関わったとき、なんらかの成長過程あるいは代謝の変化が片方あるいは両方の細胞で生じ、B に活動電位を発生させる細胞のひとつとしての A の影響力が増加する。

ある細胞が、べつの細胞が活動電位を発生するのにより多く関わる方法に関して、いちばん単純で、私をもっとも可能性があると考えている考え方は、入力側の軸索と出力側の細胞体との間でシナプスができ、接触面積を増やしていくというものである。

シナプスの可塑性が、そのシナプスの前後のニューロンの活動にだけ依存する場合には、そのシナプスは Hebb 則に従う、とよばれるのが一般的である。

実際の神経系には、Hebb 則に従わないようなシナプスもみつまっている。例えば、アメフラシ (*Aplysia*) のえら引き込み反射に関する研究でみつかったシナプスは Hebb 則とは異なる性質を示すが、これが古典的条件づけなどの基礎となる連合形成のために有効な機構になっていることが知られている [42].

海馬において、シナプスの長期増強 (long-term potentiation; LTP) という現象が知られている。シナプスの長期増強とは、ある特定の条件でシナプスの伝達効率が持続的に高くなる、という現象を指している。海馬の LTP が実際に Hebb 則に従う現象であることが、確かめられている。

海馬に損傷を受けた患者は、新しい事項の記憶ができなくなる。このことから、海馬は記憶の形成や学習といった機能をになっているとされている。そして海馬の LTP は、

海馬の学習機能の神経生理学的な基礎となる機構であると考えられている。この点に関しては、近年数多くの研究がおこなわれている。MITの Tonegawa らの研究グループは、遺伝子ターゲティング法によって $\alpha$ -カルシウムカルモジュリン依存性キナーゼ II を生成できないような突然変異マウスをつくり、そのマウスの海馬 (の CA1 とよばれる領域) が LTP を示さないこと [81] と、そのマウスが空間認知に関連する課題をうまく学習できないこと [80] とを示した。この結果は、LTP にみられるような Hebb 形のシナプスの可塑性が、学習に関する神経生理学的な機構であることの有力な証拠であると考えることができる。

海馬におけるシナプスの可塑性は、後天的な学習に関連しているが、いっぽうで、個体発生の段階での神経系の自己組織的形成においても、シナプスの可塑性が重要な役割を果たしている。視覚系の自己組織的形成において Hebb 形のシナプスが見つかっており、それが視覚系の構造の形成に関わっていることが指摘されている [79]。

## 2.4 初期視覚系

### 2.4.1 概要

視覚情報は、光のかたちで目に入射し、網膜にある視細胞によって電気的な信号に変換されて、中枢へ送られていく。解剖学的には、視覚情報は網膜から外側膝状体 (lateral geniculate nuclei; LGN) を経て大脳皮質の一次視覚野へとつながる神経線維によって運ばれる (図 2.1)。本論文では、視覚系のなかで視覚情報処理の初期段階にあたる、網膜から一次視覚野に至る部分を初期視覚系とよぶ。

### 2.4.2 網膜

網膜は眼底にあるが、発生学的には脳の一部がせりだしてきて形成されるので、神経系の一部として考えることができる。実際に、網膜には図 2.2 に示すように、光をうけとる視細胞のほかに水平細胞、双極細胞、アマクリン細胞、神経節細胞という4種類のニューロンを含む、8層の層構造がみられる。視覚情報はこれらのニューロンによって、網膜ですでに情報処理を受けている。視覚情報は主に、視細胞、双極細胞、神経節細胞という経路で伝達される。水平細胞、アマクリン細胞は、補助的なはたらきをしているとされている。

光を電気信号に変換する視細胞は、網膜のいちばん外側に層をなしている。視細胞の分布は一様ではなく、眼球の光軸上にあたる中心窩の付近で高密度であり、そこから離れるにしたがって低密度になる。

感覚系のあるニューロンが影響をうける感覚器官の領域を、その細胞の受容野という。双極細胞は視細胞からの情報をうけて応答するが、その応答のパターンから2つの種類に分類することができる。ひとつは、受容野の中心部が明るくて周辺部が暗いときに正の応答をするものである。これを、on-中心形ニューロンという。もうひとつは、逆に受容野の中心部が暗くて周辺部が明るいときに正の応答をするものである。これを、off-中心形ニューロンという。

網膜のいちばん内側に層をなしているのが、網膜神経節細胞とよばれるニューロンである。網膜にある他のニューロンは特殊なものであり、一般に活動電位を発生せずに、細胞膜の電位変化そのものをシナプスによって伝達している。けれども、網膜神経節細胞に至ってはじめて、活動電位が発生する。網膜神経節細胞もやはり、on-中心形、off-中心形の2種類のニューロンに分類される。これとは別の立場から、X形細胞もしくは小細胞、Y形細胞もしくは大細胞、W形細胞、という3種類への分類がなされることもある。小細胞は受容野が小さく、刺激に対して持続的な応答をするので、物体の形といった静的な情報処理に関与していると考えられている。大細胞は受容野が大きく、刺激に対して一過性の応答を示すので、運動情報処理に関係していると考えられている。

網膜神経節細胞の軸索は、束ねられて、視神経として眼球をでて、中枢へと向かう。小細胞、大細胞、その他、という分類によって、それぞれに属するニューロンがどこに接続しているかが異なっている。小細胞、大細胞はともに外側膝状体に接続している。少数存在する、その他のニューロンは、上丘とよばれる部位に接続していて、物体を目線で追跡する際の眼球運動や、瞳孔の光に対する反射に関係する。目からの視覚情報は、これら3つの経路によって運ばれる。このうち中枢での視覚情報処理に重要なのは小細胞、大細胞にはじまる経路であるとされている。以下ではこれら2つの経路に関して述べる。

### 2.4.3 外側膝状体

両眼を出た視神経は、視床にある外側膝状体 (LGN) にはいる。LGN は左右に一個ずつあり、それぞれが両眼からの視神経を受けているが、左にある LGN は左目の網膜の左半分と右目の網膜の左半分とから入力を受ける。同様に、右にある LGN は、両眼の網膜の右半分から入力を受ける。網膜上の視覚像は倒立像だから、LGN は両眼の視野の「反対側の半分」からの入力を受けていることになる。このような視神経の「いれかえ」は、網膜と LGN との間にある、視交叉とよばれる場所でおこなわれる。

ヒトやサルの LGN は 6 層構造をしている。そのうち内側の 2 層は、大細胞層 (magnocellular layers) とよばれ、残り 4 層は、小細胞層 (parvocellular layers) とよばれる。それぞれは、網膜神経節細胞の対応する種類の細胞からの入力を受ける。左右の目からの神経線維の入力は、混ざり合わないで各層に分離している。小細胞層がなぜ 4 層あるのかは、現在のところわかっていない。

網膜から LGN の各層への神経線維の接続は、大域的にみれば網膜での 2 次元的な位相が保たれるようなかたちになっている。

### 2.4.4 一次視覚野

LGN のニューロンからの軸索は、大脳皮質の後頭部にあたる一次視覚野へ接続している。一次視覚野はまた有線野、線条野、Brodmann の 17 野とも呼ばれ、大脳皮質の全面積の 1/3 を占めるほど広い領域である。

大脳皮質は一般に厚さが 1.5 ~ 4.5 mm であり、基本的には 6 層からなる層構造をなしている。外界からの感覚情報などは、外から数えて 4 層目 (4 層という) に入力される。また、さらに高次の中枢への出力はおもに 2, 3 層から出ていくのが普通である。1 層は線維層で、ニューロンはない。

一次視覚野の入力にあたるのは、4 層のなかでさらに細かく分けたときに  $4C\alpha$  層、 $4C\beta$  層とよばれる層である。 $4C\alpha$  層は LGN の大細胞層から、 $4C\beta$  層は小細胞層から、それぞれ入力を受けている。

まず、刺激に対する一次視覚野のニューロンの選択性という、局所的な性質について述べる。 $4C\alpha$  層では、LGN と同様に on-中心形、off-中心形のニューロンがある。けれ

ども、4 層のほかの部分では、全く性質の異なる応答を示すニューロンがある。そのニューロンは、受容野内の特定の方位の線分に対して選択的に応答を示す。このような性質を示すニューロンは、方位選択性ニューロンとよばれる。これは、Hubel と Wiesel [37] が単純形細胞と呼んだものと同じものである。

4 層からの情報は 2, 3 層に伝達され、2, 3 層からさらに 5, 6 層へ伝達される。2, 3, 5, 6 層には、線分の方位だけでなく、その運動方向に対しても選択性をもつニューロンが存在する。このようなニューロンを運動方向選択性ニューロンという。これは、Hubel と Wiesel [37] が複雑形細胞と呼んだものである。

皮質のある場所にあるニューロンの方位選択性、運動方向選択性は、皮質の厚さ方向でみると一定に保たれている。また、皮質の面方向でみると、ニューロンの方位選択性、運動方向選択性が規則的に配列しているのが観察される。このような方位選択性の規則正しい配列を、方位選択性コラムという。大脳皮質には、刺激に対する選択性がこのように規則正しくコラムをつくっている例がいくつか知られている。一次視覚野には方位選択性コラムのほかに、左右それぞれの目から入力を受ける皮質上の領域が、縞状に分離した形で存在している。この構造は眼優位性コラムとよばれる。

一次視覚野の大域的な性質をみると、LGN の場合と同様、網膜から一次視覚野への神経線維の接続は、大域的にみれば網膜での 2 次元的な位相が一次視覚野上で保たれるようなかたちになっている (図 2.3)。網膜から一次視覚野への神経線維の接続が、位相を保ったかたちになっている、すなわち位相保存写像を形成していることを、網膜部位再現あるいは retinotopy という。

さらに、網膜上で視細胞の多い中心窩付近の領域は、一次視覚野上では広い面積に対応する。反対に、周辺視にあたる領域に対応する一次視覚野上の領域は、狭くなっている。一次視覚野上で単位視角が対応する長さを調べると、図 2.4 に示すように、視線方向を基準とした視角に対してほぼ反比例の関係にあるのが観察される [12]。いっぽう、網膜上での視細胞の分布をみると、明所視ではたらく杆体という種類の視細胞は、中心窩付近ではおおざっぱにいて視角の 2 乗に反比例するような分布を示す (図 2.5)。このことは、視覚刺激を視点を中心にして一様に拡大、縮小しても網膜上での分解能が変化しない、という都合のよい性質を与えていると考えることができる。

網膜での視細胞の分布の密度は、網膜上の単位面積での視覚刺激の強さをあらわすと考えることができるから、おおまかにいって網膜から一次視覚野への写像は、一次視覚野での単位面積あたりの視覚刺激を均一にするように形成されているということが出来る。これは、いかえれば網膜での単位面積に対応する一次視覚野上の面積すなわち写像の拡大率と、そこでの刺激の強さとが比例関係にある、ということでもあるので、この関係のことを、比例則とよぶ。

また、網膜から一次視覚野への写像は、よい近似で等角写像とみなせることも報告されている。Landau と Schwarz [55] によれば、網膜から一次視覚野への写像は、大域的にみれば  $w = \log z$  という等角写像がよい近似になっている。この式は、さきに述べた比例則とも両立するものである。

一次視覚野での網膜部位再現は、日露戦争の軍医であった井上により、一次視覚野への銃創と視野欠損との対応を調べることによって 1904 年にはじめて報告されたとされている。視覚系での網膜部位再現は一次視覚野だけでなく、現在ではほかのより高次の領野でも知られている。けれども、高次の領野ほど網膜部位再現は明確でなくなり、各ニューロンはより高次の情報に選択性をもつようになる傾向がある。

## 2.5 考察

神経系の自己組織的形成に関する問題は、神経系のふるまいの変化すなわち神経系の可塑性を取り扱うという点で、神経系の学習に関する問題と共通点をもっている。神経系の学習あるいは自己組織的形成に関する研究の、現在での主な流れとしては、以下の 4 つを挙げることができる [10]: (1) 古典的条件づけや馴化などの、単純な可塑性の神経生理学的な機構を調べるもの。(2) 海馬などの側頭葉を解剖学的、生理学的に調べるもの。(3) 発生段階での視覚系の形成を調べるもの。(4) 遺伝子と神経系の形成との関係を調べるもの。

本論文での議論は、おもに発生段階での視覚系の形成に関係している。視覚系を研究の対象とする利点はいくつかある。生理学的には、実験条件のコントロールが容易であるために実験がやりやすいことがあげられる。この利点のおかげで、多くの知見が蓄積されており、それらが利用できる。また、視覚系のニューロンは脳の他の部位における



ニューロンと基本的に同じである。視覚野に限らず、大脳皮質の各領野は、その機能にかかわらず基本的に同じ構造をもっている。事実、網膜から LGN へ向かうべき視覚神経を聴覚野へつなげる内側膝状体へ外科的に導いた研究 [76, 83] では、聴覚野においても通常の視覚野と同様に、視覚刺激に対する方位選択性細胞や位相保存写像が形成されたことが報告されている。

工学的な立場から取り扱うことを考えた場合にも、視覚系を研究の対象とする利点はある。とくに初期視覚系では、取り扱われる視覚情報は幾何学的に単純なものであると考えることができる。それを取り扱う視覚系自身も、大規模な処理系でありながら構造的、機能的に一樣であるとみなしてよいような性質をもっている。これらのことが、視覚系は工学的な取り扱いに向いている対象だということの理由である。

この章で述べた知見を、本論文での以下の議論に対応させて、単純化したかたちでまとめおく。視覚系は、ニューロンを構成単位とする神経回路網として捉えることができる。初期視覚系は大域的にみると、小細胞経路、大細胞経路のそれぞれが網膜にはじまる何層もの層状構造を形づくっていて、網膜からの視覚情報はこれらの層を順に伝達されていくようになっている。その伝達をになう視覚神経線維の接続は、刺激の頻度による重みづけを伴う位相保存写像といえるものである。また細かくみると、各層には on-中心形や off-中心形、方位選択性、運動方向選択性といった、刺激に対する選択性をもったニューロンが存在する。はじめの方の層には、比較的単純な選択性をもつニューロンがあり、あとの方の層ほど、複雑な刺激に対して選択性をもつニューロンが表れてくる。

神経系で神経線維がどのように走り、どの部分からどの部分への接続があるかということ詳しく調べると、ほとんどの場合に双方向性の接続が観察される。例えば、LGN は一次視覚野へ神経線維を送っていることは既に述べたが、実は一次視覚野から LGN へ、という逆の向きへも神経線維の接続がある。この接続は、LGN から一次視覚野への接続と比較すると、線維の数にしておよそ 10 倍と見積もられるほど多いものである。

感覚系におけるこのような逆行性の接続がどのような機能を果たしているか、ということに関しては、様々な見解が述べられている。例えば、高等動物では、視床は嗅覚以外のほとんどすべての感覚入力で大脳皮質へ中継しているが、各感覚野から視床

への逆行性の情報によって、ある種の「センサフュージョン」のはたらきを視床がはたしているのではないか、という考え [70] も提案されている。また、感覚野への情報の流れでみられる「側抑制」とよばれる現象が、実際の抑制性ニューロンの接続パターンと対応していない、という観察から、逆行性のフィードバックによって等価的に抑制が実現されている、とする考え [72] も提案されている。

このように、感覚系における逆行性の接続のはたらきについては、多くの仮説が提案されているものの、現在までにこれといった定説はない。したがって本論文では、基本的に順方向の結合だけを考え、逆行性の接続については取り扱わない。

## 第 2 章 図表

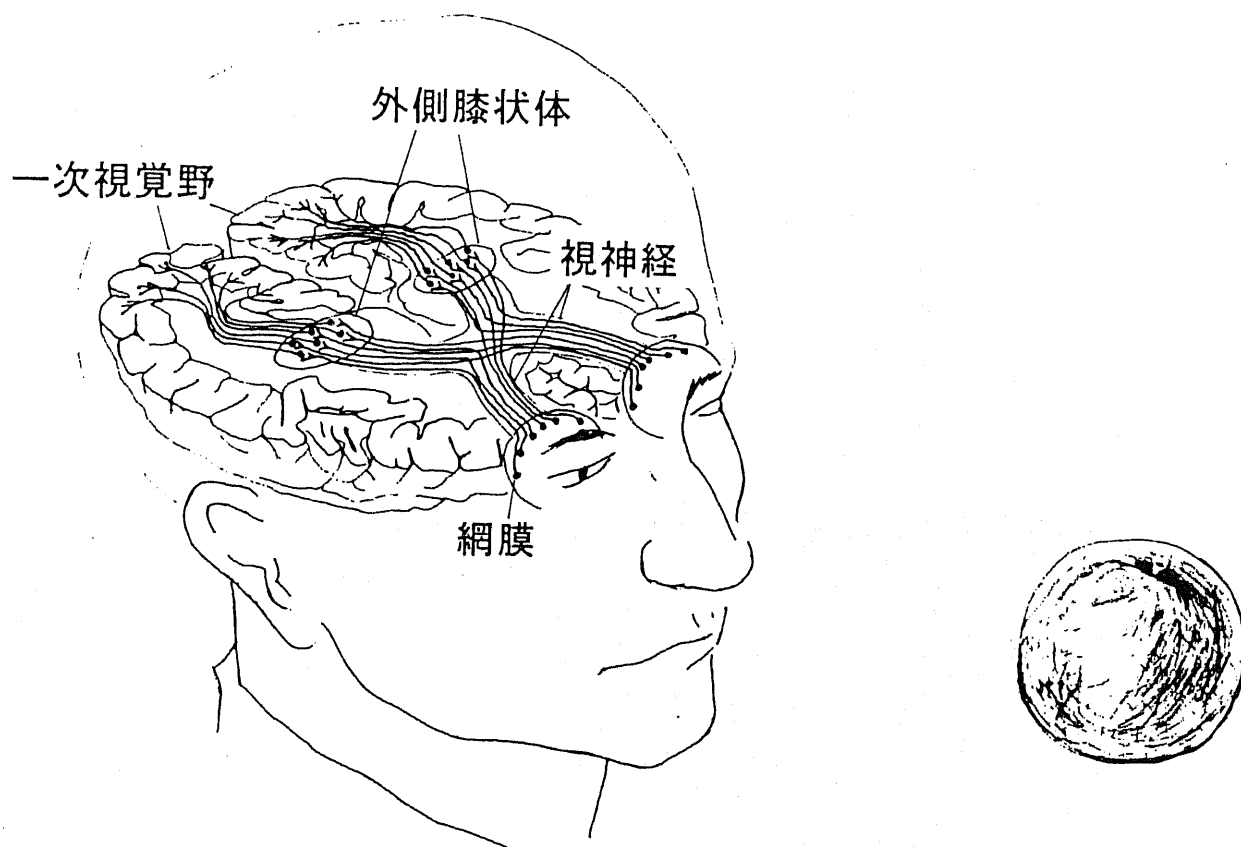


図 2.1 初期視覚系の構成

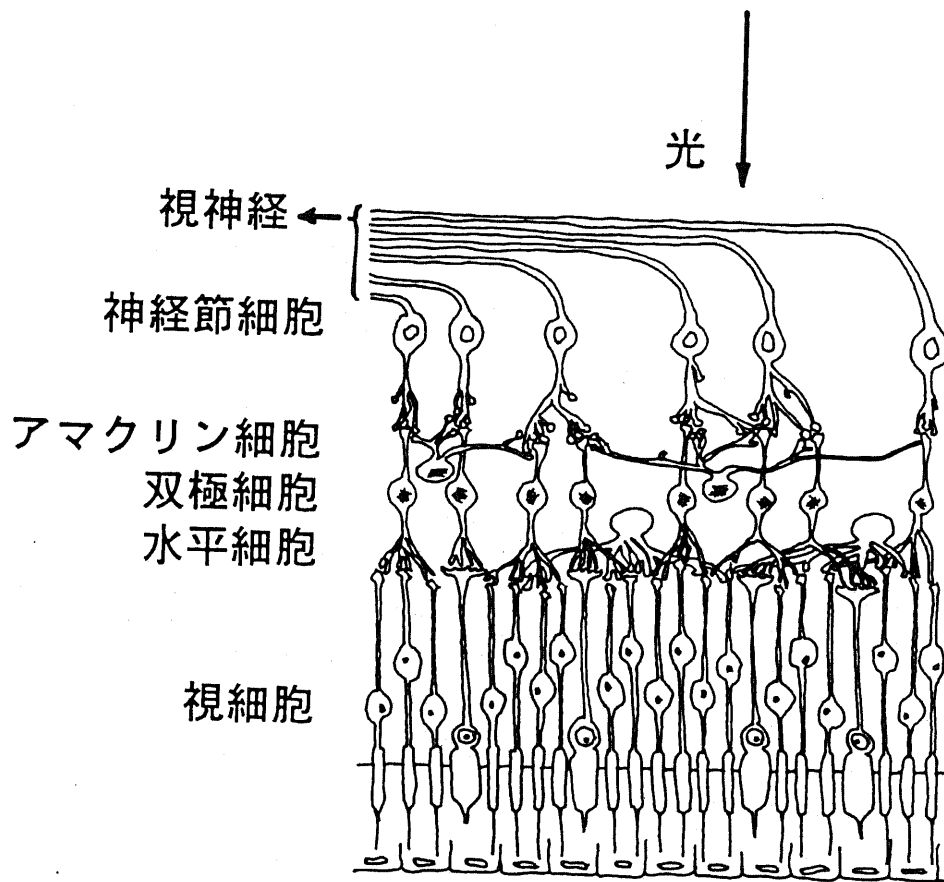


図 2.2 網膜の構造

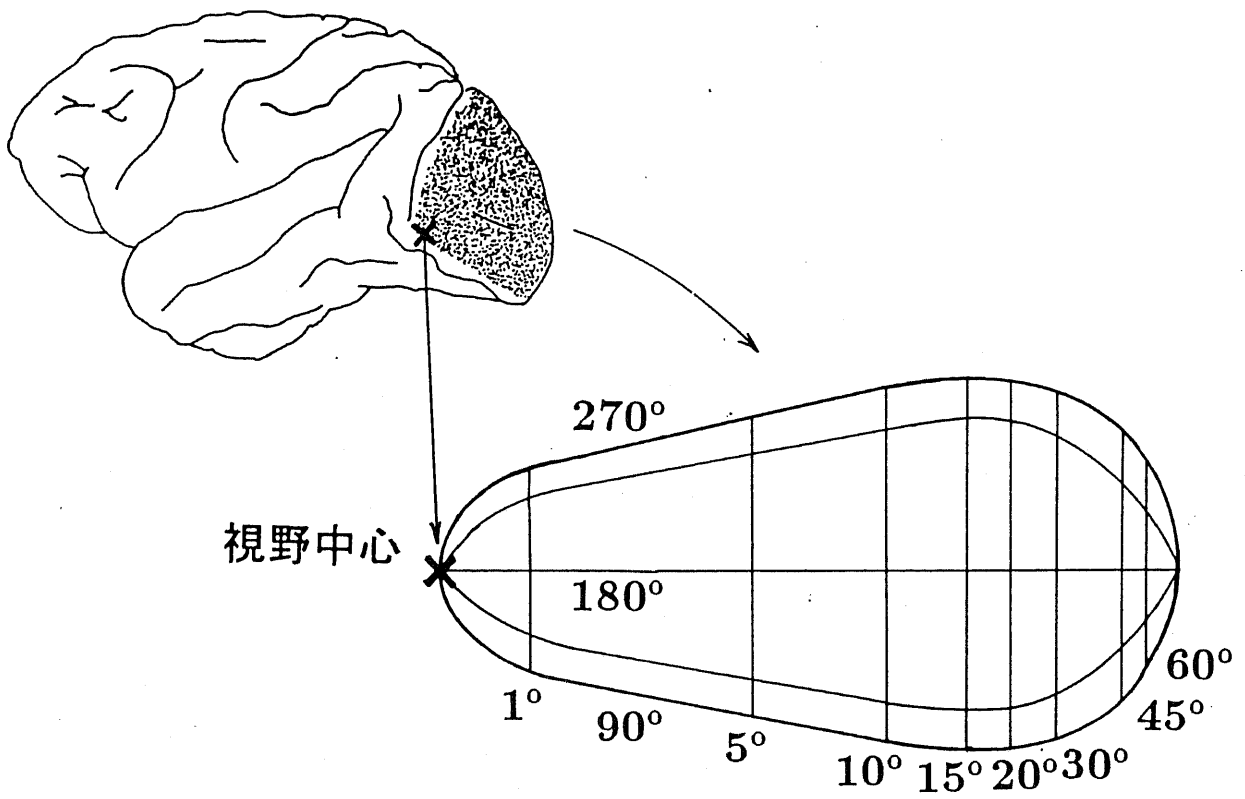


図 2.3 一次視覚野上での網膜部位の再現

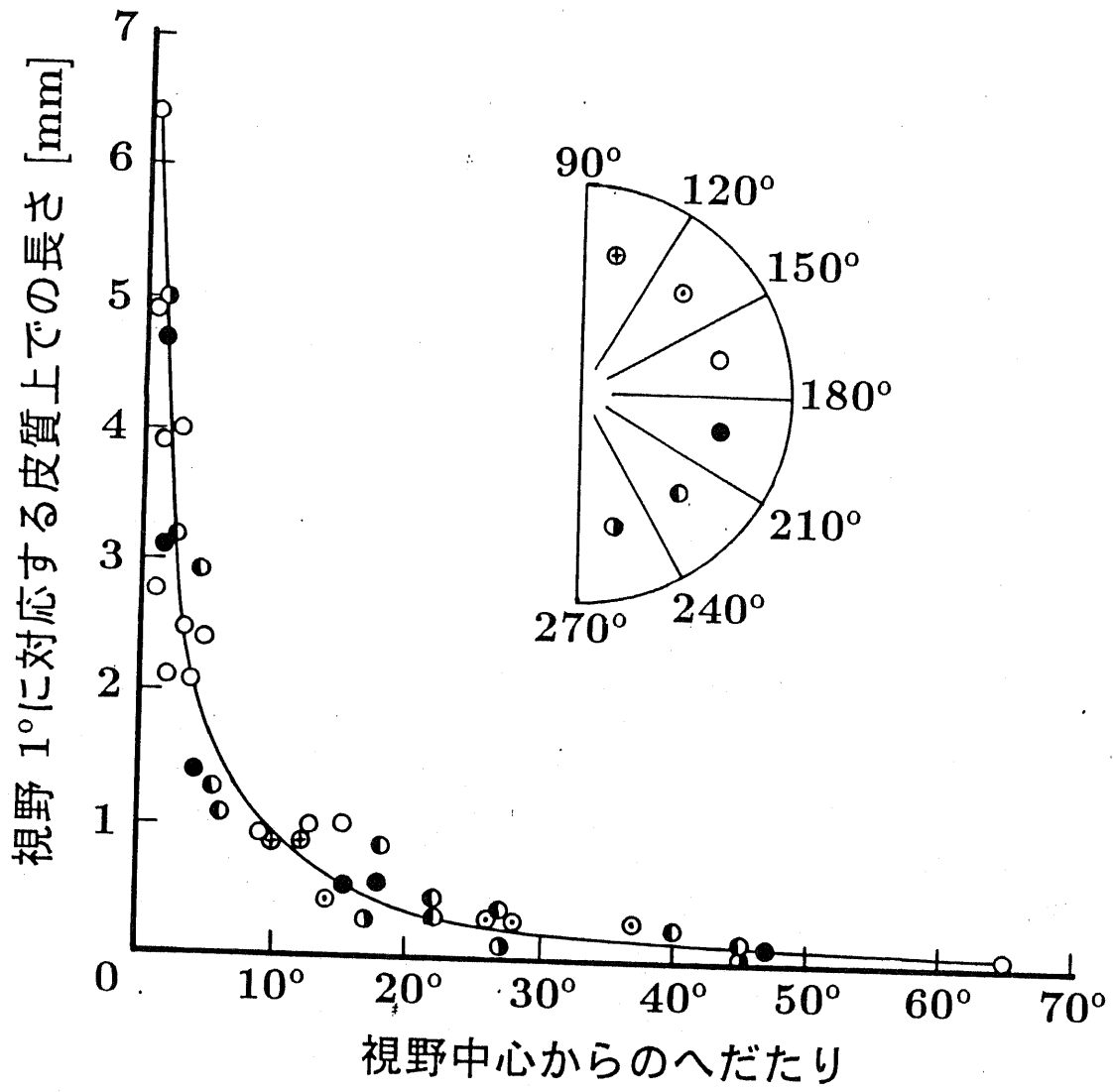


図 2.4 一次視覚野上での単位視角に対応する長さ と 視線方向からの視角との関係

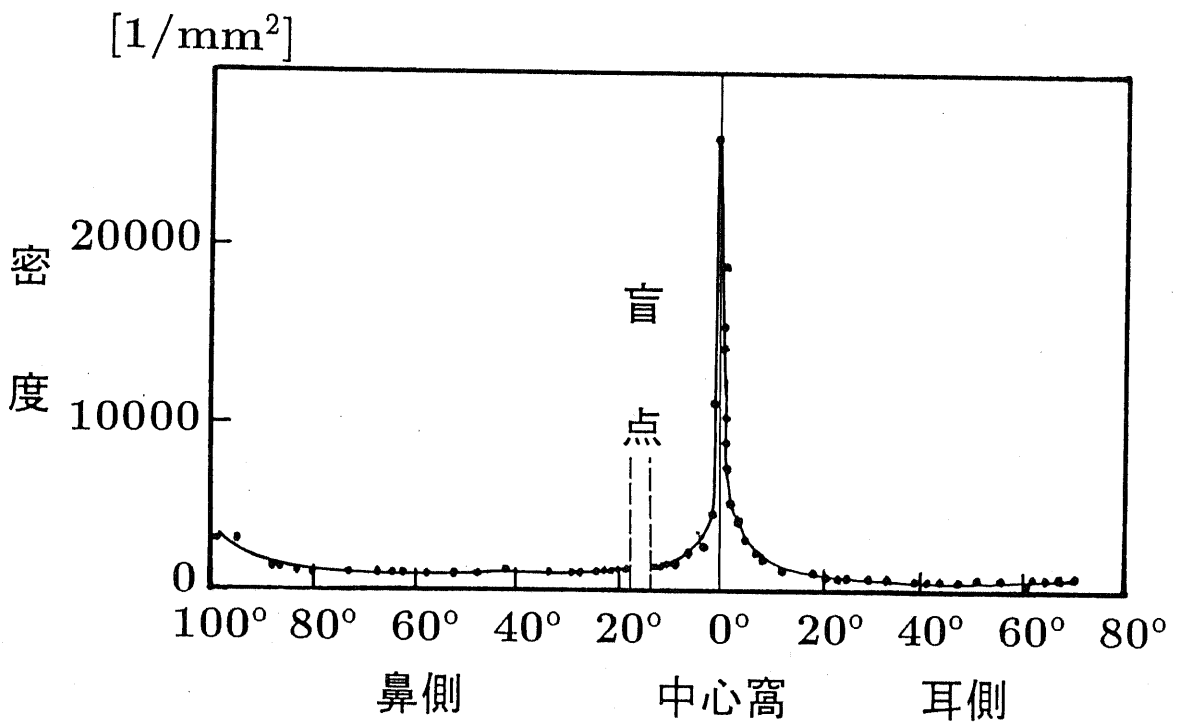


図 2.5 網膜上での杆体の分布



## 第 3 章

# 神経回路網理論に関する基礎的事項

すべて男も女も、わるものは、わづかにしれる  
かたの事を、のこりなく見せつくさむと思へる  
こそ、いとおしけれ ([71])

### 3.1 本章の目的

神経系を、単独では単純な動作をする形式ニューロンが相互に結合したものと取り扱おうというのが、神経回路網理論の立場である。

実際の神経系の構成要素であるニューロンに対応する、神経回路網の構成要素は、形式ニューロンである。学習および自己組織化のための規則として、本論文では、Hebb 則を具体的にアルゴリズムの形で与えたものを考える。また、実際の神経系が層構造をなしていることに対応する、神経回路網理論での道具立てとして、神経場の考え方を導入する。本章では、これらのことについて概説する。

### 3.2 形式ニューロン

神経回路網の構成要素として、実際のニューロンを形式化、単純化したものを考える。これを形式ニューロンという。本論文では、混同のおそれがない限り、形式ニューロンのことも単にニューロンということにする。

形式ニューロンは、多入力 1 出力の演算素子である。入出力関係は、一般にはどんなものを考えてもよいが、各入力にそれぞれ重みを掛けて和をとり、それをある関数（出力関数とよばれる）に通したものが出力になる、という演算を考えるのが普通である（図 3.1）。

これを数学的に表現してみよう。ニューロンは  $N$  個の入力をもつとする。それぞれの入力の値を  $x_i$  ( $i = 1, \dots, N$ ) であらわす。また、 $N$  個の入力それぞれに割り当てられる重みを結合荷重といい、 $w_i$  ( $i = 1, \dots, N$ ) であらわす。出力関数を  $f$  であらわすと、ニューロンの出力  $y$  は、以下のようにあらわされる。

$$y = f\left(\sum_{i=1}^N w_i x_i\right) \quad (3.1)$$

あるいは、 $\boldsymbol{x} = (x_1, \dots, x_N)$ ,  $\boldsymbol{w} = (w_1, \dots, w_N)$  をそれぞれ入力、結合荷重を成分とするベクトルとすると、以下のような簡便な表記でかける。

$$y = f(\boldsymbol{w} \cdot \boldsymbol{x}) \quad (3.2)$$

ここで、 $\cdot$  はベクトルの内積をあらわす。

内積  $w \cdot x$  は、ある意味でベクトル  $w$  と  $x$  との類似度をあらわしていると考えることができるから、ニューロンの入出力関係として、以下のように入力と結合荷重ベクトルとの差の関数として出力が決まるようなものを考えることもできる。

$$y = f(w - x) \quad (3.3)$$

この形の入出力関係をもつ形式ニューロンも、よく使われる。

本論文では、以上のような入出力関係をもつ形式ニューロンを取り扱う。非常におおざっぱに言えば、入力を線形加算することは各入力に由来する細胞膜の電位が線形に加算されることに対応し、出力関数は活動電位を発生するような細胞膜の能動的な性質を反映している。また、結合荷重は、シナプスの伝達効率に対応している。

このようなニューロンが多数、相互に結合することで、神経回路網が形成される。簡単な神経回路網の例として、ニューロンの2つの集団  $L, M$  があり、集団  $L$  のニューロンのそれぞれからから集団  $M$  のそれぞれのニューロンへ向かう結合があるようなものを考える。集団  $L, M$  にはそれぞれ  $L$  個、 $M$  個のニューロンがあるとしよう。集団  $L$  に属する  $l$  番目のニューロンから、集団  $M$  に属する  $m$  番目のニューロンへの結合荷重を  $w_{ml}$  とあらわし、ニューロン  $l$  の出力を  $x_l$ 、ニューロン  $m$  の出力を  $y_m$  とかくと、この神経回路網は以下のような式で表現することができる。

$$y_m = f\left(\sum_{l=1}^L w_{ml}x_l\right) \quad (3.4)$$

あるいは、 $y$  を、 $y_m$  を成分とするベクトル、 $x$  を、 $x_l$  を成分とするベクトルとし、 $W$  を、 $w_{ml}$  を要素とする行列とすると、式 (3.4) はより簡便なベクトル形式でかける。

$$y = f(Wx) \quad (3.5)$$

ここで、ベクトル値関数  $f(z)$  は、 $z$  の各成分にスカラー関数である出力関数  $f(z)$  を作用させる関数であるとする。

McCulloch と Pitts [68] が考えた形式ニューロンは、おおざっぱに言えばここで説明した形式ニューロンで出力関数がステップ関数になっているものであるといてよい(厳密には、抑制性の入力への取扱いが異なる)。いいかえれば、McCulloch-Pitts ニューロンは論理素子であるということになる。彼らは、この形式ニューロンを使えばいかなる論理関数も実現できること、すなわち論理関数を実現するための構成要素としてこの形式ニューロンが完全性をもつことを示した。

その後、実際のニューロンが活動電位の時間的な頻度で情報を表現している、という考え方のほうが、広く受け入れられるようになっていく。それに対応して、形式ニューロンを考えると、ステップ形でなく出力値として連続値をとるような出力関数、とくに S 字形の出力関数を使うのが普通になっている。S 字形の出力関数をもつ形式ニューロンについても、原理的には、それを組み合わせて任意の関数を任意の精度で近似できる、という意味での完全性が示されている [7, 8, 20-23, 26, 30, 35, 36, 39, 40, 53, 54]。

神経回路網の「動作」を議論するためには、時間関数として入力を与えられた場合のニューロンの動作を表現する必要がある。通常、ニューロンの動作を微分方程式や差分方程式で記述することが行われる。一定の入力に対するニューロンの状態変化が、ある一定の安定平衡点に指数的に緩和するようなものであるときには、入力の時間的な変動の時間スケールが、ニューロンの動作に関する時間スケールよりも十分に長ければ、ほとんどの時刻で、ニューロンの状態は入力できまる安定平衡点にあると考えてよい。実際のニューロンでは、カオス的な動作が観察されることから、形式ニューロンの動作としても上で述べたような緩和形と違う、カオス的な動作規則が採用される場合もある [2]。けれども本論文では、形式ニューロンとしては単純なものを考えて回路網の水準での考察をしたいので、このような複雑な形式ニューロンは考えない。また、神経回路網の「動作」ではなく、神経回路網の自己組織的形成という、もっと時間スケールの長い現象を対象にするので、ニューロンの時間的なダイナミクスは議論しない。

### 3.3 Hebb 則

2.2 で述べたように、生体の神経系では、その動作の変化はおもにシナプスの可塑性にもとづくものであると考えられる。神経回路網理論の立場では、回路網の性質の変化は、シナプスに対応する結合荷重が変化することによって取り扱われる。

神経回路網の結合荷重を問題に適したやりかたで変化させるための規則は、神経回路網の学習則とよばれる。神経回路網の学習則として、シナプスの可塑性における Hebb 則に対応するものを考えることができる。2.2 で述べたように、シナプスの可塑性について Hebb が提案しているのは、明らかに原理であってアルゴリズムや機構ではない [10]。けれども、神経回路網理論での学習則としては、アルゴリズムの形で Hebb 則を表現する必要がある。

Hebb 則の具体的な表現は一通りに限られない。二つのニューロンの間の結合が、その二つのニューロンの活動をそろえるように変化する、という Hebb 則の基本的な考え方にしたがえば、ニューロンの入出力関係が式 (3.2) で与えられるときに、結合荷重ベクトル  $w$  を次の式によって変化させるのが、Hebb 則の基本形であるということが出来る。

$$\Delta w = \varepsilon y x \quad (3.6)$$

ここで  $\varepsilon$  は小さい正数とする。

この形のままで、結合荷重ベクトルが学習とともに発散してしまうことがあるので、結合荷重ベクトルとして可能な値に制約を設けたり、結合荷重ベクトルを正規化したり、といった工夫とともに使われるのが普通である。例えば、von der Malsburg [94] は、結合荷重ベクトルの成分の総和が一定であるという条件を使っている。式 (3.6) に指数形の減衰項を付加した、次のような形もよく使われる。

$$\Delta w = -\varepsilon w + \varepsilon y x \quad (3.7)$$

ニューロンの入出力関係が式 (3.2) で与えられるときには、基本形が次式のような形をとる学習則も Hebb 則としてよく使われる。

$$\Delta w = \varepsilon y(x - w) \quad (3.8)$$

### 3.4 神経場

第 2 章でみたように、視覚系をはじめとして感覚系では、感覚受容器から大脳皮質にある感覚野までの経路は、ニューロンが層をなしているような構造をしている。しかも大域的には、視覚系における網膜部位再現のような、各層での位相を保つような形で各層間の結合がみられる。神経場の考え方は、このような構造的特徴を形式的に取り扱うためのものである。

ニューロンの層を一層分取り出して考えよう。この層は典型的には 2 次元のシートだと考えてよいが、一般には何次元でもよい。この層にはたくさんのニューロンが一様に存在しているとしよう。非常にたくさんのニューロンを考えるとときには、ひとつひとつのニューロンははっきりとは見えなくなって、層の各点がひとつのニューロンの性質を示すような、ニューロンの連続的なシートのように捉えることができる。このシートのことを、(2 次元の) 神経場とよぶ。

神経場による単純な神経回路網の例を考えよう。いま、 $L$ 、 $M$  という 2 つの神経場があり、さらに  $L$  から  $M$  への一方向性の結合があるような神経回路網を考える (図 3.2)。 $L$ 、 $M$  のそれぞれのニューロンは、それぞれの神経場での座標を指定することで特定することができる。神経場  $M$  に属する、位置  $y$  にあるニューロンを考えよう。このニューロンは、神経場  $L$  の位置  $x$  にあるニューロンからの出力信号  $s(x)$  を、結合荷重  $w(y, x)$  でうけとる。したがってこのニューロンの出力  $s(y)$  は、式 (3.5) を自然に拡張した形をした、以下の式で表されることになる。

$$s(y) = f \left( \int_{x \in L} w(y, x) s(x) dx \right) \quad (3.9)$$

神経場に対しても、Hebb 則は自然に拡張される。

### 3.5 考察

神経回路網の学習則として古典的なものに、パーセプトロン学習則がある [69]。パーセプトロン学習則は、学習の収束が理論的に保証されている、数少ない学習則のひとつである。けれどもこの学習則は、線形分離可能課題という、狭いクラスの問題に対してしか適用できない。

実用的な神経回路網の学習則としては、誤差逆伝搬学習則 (back propagation) が有名である。誤差逆伝搬学習則の歴史は古く、込み入っており、多くの研究者が独立に、同様の学習則を提案してきた [31]。けれども、誤差逆伝搬学習則を実用的な学習則として整備し、広く浸透させたのは、Rumelhart ら PDP グループ [77] の功績である。誤差逆伝搬学習則は、学習の収束の理論的な保証こそないものの、多くの問題に適用されてよい結果をもたらしてきた。そのため、実用的な学習則として現在広く使われている。

神経系の学習のモデルとして誤差逆伝搬学習則を考えようとする、別の問題が生じる。生体の神経系では、あることを学習するためにそれぞれのシナプスが伝達効率を変化させるときには、そのシナプスの場所で手にはいる局所的な情報しか使うことができない。また、すでに述べたようにシナプスの伝達は一方向性だから、誤差逆伝搬学習則でおこなわれるような、情報の逆向きの伝達を実現するには、それに対応した、逆方向の神経線維もなければならぬ。実際、うまく神経回路網を構成して、その上でこれらの要件をみたすように誤差逆伝搬学習則を実現することは可能だけれども [29]、少なくとも神経系の学習のモデルとしては、誤差逆伝搬学習則は現実的ではない。

これに対して、Hebb 則は局所的な情報だけを使って学習を行なうので、神経系の学習のモデルとしては優れている。定性的には、実際に Hebb 形の可塑性を示すシナプスがみついているのも、Hebb 則の有利な点である。工学的な立場から考えても、局所的な情報だけを使う適応的システムは、ハードウェアでの実現がしやすい、という利点がある。大規模なシステムになればなるほど、システムの各構成要素間の通信すなわち情報のやりとりが、大きな問題になってくるからである。それだけに、このような「自律分散システム」が大域的にどのようなふるまいをするのか、ということ、実験的にであれ、理論的にであれ調べるのが重要である。

### 第 3 章 図表



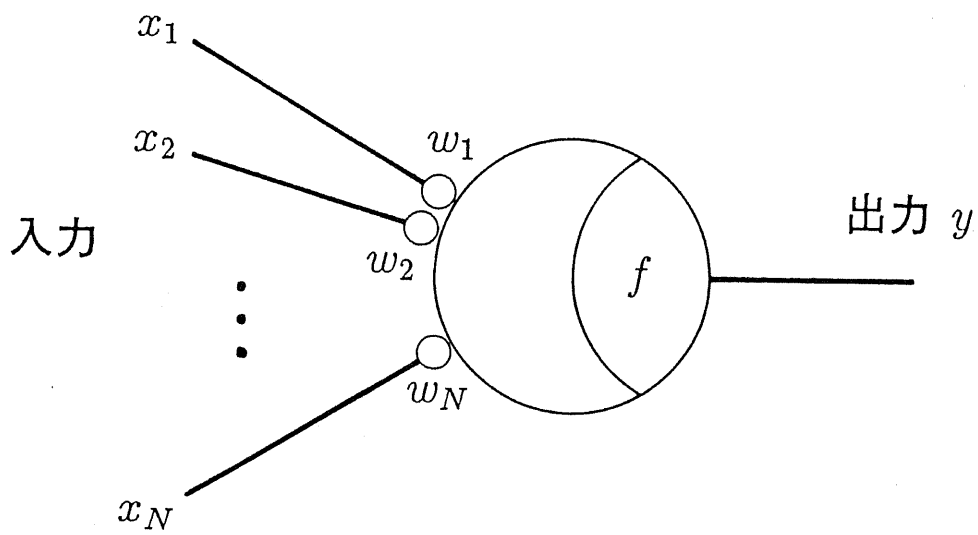


図 3.1 形式ニューロン

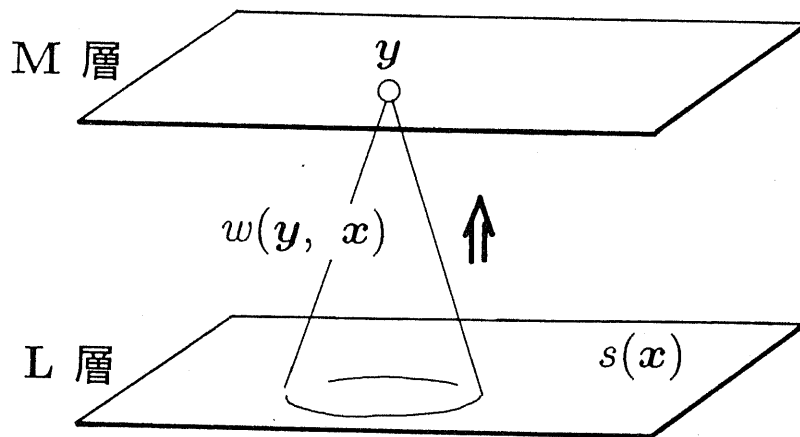


図 3.2 神経場による単純な神経回路網

## 第 4 章

# 構造の大域的な特徴の自己組織的形成 に関するモデル

— Dieß sind die Bedingungen des Problems.

Hic Rhodus, hic salta. ([67])

## 4.1 本章の目的

ヘルシンキ工科大学の Kohonen によって提案された自己組織化特徴写像モデル (self organizing feature map; 以下では Kohonen のモデルとする) は、神経系の自己組織的形成に関するモデルとしては、例えば視覚系にみられる、位相保存写像の形成や、写像の拡大率が刺激の頻度に対応しているといった大域的な特徴と同様な性質をもったモデルである。いっぽう、神経系の自己組織的形成のモデルという立場から離れて、純粋に工学的なシステムとしてみても、適応的信号処理をおこなうシステムであると考えることができる。事実、Kohonen のモデルは適応的信号処理システムとして、様々な問題に実際に適用されている。

実際の問題への応用においても、入力信号空間からの位相保存写像を自己組織的に形成するという性質や、入力信号の頻度に対応する写像の拡大率を実現する性質が、有効に利用されている。けれども、これらの性質は、計算機シミュレーションなどによって実験的に確かめられているにすぎない。Kohonen のモデルによって位相保存写像が形成されるためにはどのような条件が必要なのか、あるいは、Kohonen のモデルによって実現される写像の拡大率が入力信号の頻度とどのような関係にあるのか、といったことについて、実験的にアルゴリズムの性質を検討した研究はいくつかある [1, 13, 63] もの、理論的な検討はほとんど行われていない。本章では、これらの問題を議論する。

適応的信号処理アルゴリズムとして Kohonen のモデルをみると、Kohonen のモデルは、競合学習とよばれる方式と、協調学習とよばれる方式とを組み合わせたアルゴリズムであるということができる。競合学習という方式は、適応的信号処理の分野では一般的な方式であり、最近接符号化 (nearest-neighbor coding) などで広く使われている。これと協調学習とを組み合わせたところが、Kohonen のモデルの特徴であるということができる。

神経系の構造の大域的な特徴が自己組織的に形成されることに関するモデルとしてみるならば、Kohonen のモデルは2層フィードフォワード形の層状神経回路網であるということができる。ただし、信号が入力されるほうの層が神経場として取り扱われ、出力層とのあいだの結合については、結合を平均化したもので代表するような近似をした

形となる。また、競合学習、協調学習を実現するために、Kohonen のモデルは出力層内で相互結合をもっている。

このように Kohonen のモデルは、第 3 章で説明したような単純な神経回路網というよりも、簡単な構造とそれによる機能とを、神経回路網の水準よりもうひとつ上位の水準で形式化したうえで定義されているとみることができる。

本章ではまず、Kohonen のモデルの概略を説明し、神経系との関係について議論する。つぎに、Kohonen のモデルによって実現される写像の拡大率が入力信号の頻度とどのような関係にあるかを、アルゴリズムの細かい設定との関係で定量的に明らかにする。Kohonen のモデルによって位相保存写像が形成されるための条件に関しても議論する。この問題を取り扱うのは困難なので、問題が比較的やさしい 1 次元の場合をおもに取り扱う。一般の次元では、粗視化した形でこの問題に関する議論を試みる。

## 4.2 Kohonen のモデル

### 4.2.1 基本的な競合学習

本節ではまず、Kohonen のモデルの基礎となる基本的な競合学習について、適応的信号処理アルゴリズムとしての立場から説明する。神経場の枠組みとの関係については、あとで議論する。

$n$  次元の信号空間  $S$  を考える。入力信号  $x \in S$  は、確率密度分布  $p(x)$  に従って独立に生起するとする。これとは別個に、 $N$  個のニューロンからなる集団  $F$  を用意する。 $i$  番目のニューロンは、 $S$  に値をとる参照ベクトル  $w_i$ , ( $i = 1, \dots, N$ ) をもっている。

基本的な競合学習では、以下の手順で学習が行われる (図 4.1): 入力信号  $x$  が発生したとき、 $x$  に最も近い値をもつ参照ベクトル  $w_{i(x)}$  を探す。 $w_{i(x)}$  から  $x$  へ向かう単位ベクトルを  $e_{i(x)}$  とし、 $w_{i(x)}$  と  $x$  との距離が  $S$  での適当なノルム  $D$  によって  $D(x - w_{i(x)})$  とかけられるとすると、参照ベクトル  $w_{i(x)}$  に対して次式にしたがって学習が行われる。

$$\Delta w_{i(x)} = \varepsilon D(x - w_{i(x)}) e_{i(x)} \quad (4.1)$$

ただし,  $\varepsilon$  は小さい正数である. また, 本論文で「ノルム  $D$ 」というときには, 数学的に厳密な意味での「ノルム」でなく, 単に  $D(x-y)$  によって  $x, y$  の「距離」を誘導するもの, という意味を表しているものとする. 確率密度分布  $p(x)$  にしたがう入力信号  $x$  の系列を考え, それぞれの入力信号  $x$  に対してこの操作を繰り返すことによって, 学習が行われる. 式 (4.1) を, 基本的な競合学習に関する学習方程式という.

ここで述べた基本的な競合学習のアルゴリズムを適応的信号処理アルゴリズムとしてみると, 適応的ベクトル量子化アルゴリズムの一種であると考えることができる. 実際, 基本的な競合学習のアルゴリズムは, Lloyd が考案し, Linde, Buzo, Gray によって整備されたいわゆる LBG アルゴリズム [56] や,  $k$ -平均法 [15] といったベクトル量子化アルゴリズムと類似している.

#### 4.2.2 協調学習

Kohonen は, 前節での基本的な競合学習アルゴリズムに, 「近傍系」による協調学習を取り入れた形の学習則を提案している. この学習則は Kohonen のモデルの上で実現される. 本節では, Kohonen のモデルで取り入れられている協調学習について述べる.

Kohonen のモデルでは, ニューロンの集団  $F$  のなかで, 入力信号と同じ  $n$  次元の格子状に各ニューロンが規則正しく整列していると考え. ニューロン  $i$  の近傍  $N_i$  は, この格子状の配列でみてニューロン  $i$  の近くにあるようなニューロンの集合である, とする. また, 近傍の「大きさ」というものを考え, これを  $m$  であらわす.  $m=0$  なら  $N_i = \{i\}$  である. 近傍の大きさ  $m$  が大きいほど,  $N_i$  にはニューロン  $i$  から遠いニューロンも含まれるようになる.  $n=2$  の場合のニューロンの格子状の配列, 近傍の例を図 4.2 に示す.

入力信号  $x$  が発生したとき, 前節の基本的な競合学習では入力に最も近い参照ベクトル値をもつニューロン  $i(x)$  だけが学習したが, Kohonen のモデルでは, ニューロン  $i(x)$  の近傍  $N_{i(x)}$  に含まれるニューロンすべてが学習する (図 4.3).

$$\Delta w_j = \varepsilon D(x - w_j) e_j \quad (j \in N_{i(x)}) \quad (4.2)$$

$$= 0 \quad (j \notin N_{i(x)}) \quad (4.3)$$

式(4.3)を, Kohonen のモデルに関する学習方程式という.

#### 4.2.3 神経場モデルおよび神経系との関連

前節では, Kohonen のモデルの定義を天下りに与えた. 本小節では, Kohonen のモデルと, 神経場モデルとの関連について議論する.

ある神経場からひとつの出力ニューロンへの結合を考える.  $S$  を神経場とする. 神経場  $S$  の点  $x$  から出力ニューロンへの結合荷重が  $w(x)$  であるとする. 神経場  $S$  の点  $x$  からの出力を  $s(x)$  で表わすと, 出力ニューロンの出力値  $y$  は以下の式で与えられる.

$$y = f(z) \quad (4.4)$$

$$z = \int_{x \in S} w(x)s(x) dx \quad (4.5)$$

いま, 神経場  $S$  から出力ニューロンへの結合荷重の空間分布  $w(x)$  が, その重心  $\tilde{w}$  を中心にして, ほぼ点対称な「さかさつりがね形」をしているものとする. また, 出力ニューロンへ入力されるのは, 神経場  $S$  の各点からの出力のパターン  $s(x)$  であるが, 出現するどの出力パターンも, その重心  $\tilde{s}$  を中心にして, 点対称な「さかさつりがね形」をしているとする. 出力パターンの空間分布の形は, 出力パターンが現れる場所にはよらずにいつも同じであるとする. すると, 出力ニューロンの出力値  $y$  は, ほぼ  $\tilde{w}$  と  $\tilde{s}$  との位置関係だけで決まり, 等価的な出力関数  $\tilde{f}$  を使って

$$y = \tilde{f}(\tilde{s} - \tilde{w}) \quad (4.6)$$

とかくことができる.

次に, このニューロンに関する Hebb 則を考える. 議論を簡単にするため, この出力ニューロンが結合荷重  $w(x)$  をもっているとき, このニューロンに  $s(x)$  であらわされる入力パターンが与えられ, それに対するこのニューロンの出力値  $y$  が 1 であったとしよう.

このとき、(減衰項つきの) Hebb 則は次の式のようにあらわされる。

$$\Delta w(\mathbf{x}) = -\varepsilon w(\mathbf{x}) + \varepsilon \mathbf{1} \cdot s(\mathbf{x}) \quad (4.7)$$

この式の両辺を神経場  $S$  にわたって積分し、それぞれの重心をとると、

$$\begin{aligned} \Delta \bar{w} &= -\varepsilon \bar{w} + \varepsilon \bar{s} \\ &= \varepsilon (\bar{s} - \bar{w}) \end{aligned} \quad (4.8)$$

となる。 $\bar{w}$  は「平均の結合荷重」であり、 $\bar{s}$  は「入力パターンの重心の位置」であると考えることができるから、式 (4.8) は、平均の結合荷重が、入力パターンの重心の位置に対して式 (3.8) であらわされるような Hebb 則にしたがって学習をすることを表している。

Kohonen のモデルでは、これを単純化して、出力ニューロンは平均の結合荷重  $\bar{w}$  を  $S$  に値をもつ参照ベクトル  $w$  として表現し、保持しているものとし、入力パターンの重心の位置  $\bar{s}$  が入力信号  $x$  を表しているともみなして、入力信号  $x$  と参照ベクトル  $w$  との距離が近いほど大きい値を出力するように動作すると考えるのである。また、Kohonen のモデルでの Hebb 則は、式 (4.8) に対応して、参照ベクトル  $w$  を入力信号  $x$  のほうへすこし移動させる、という形をとる。参照ベクトルの移動の大きさは、そのニューロンの出力値に比例し、参照ベクトル  $w$  と入力信号  $x$  との距離に依存した大きさである。

Kohonen のモデルでは、出力ニューロンは  $N$  個ある。 $N$  個の出力ニューロンは、それぞれ手持ちの参照ベクトルと、入力信号との距離を調べ、その距離に応じて、近ければ大きい値を、遠ければ小さい値を、まず出力する。つぎに、手持ちの参照ベクトルが入力信号にもっとも近いようなニューロンを、 $N$  個のニューロンのなかから選び出すという作業が行われる。

基本的な競合学習では、この作業のために、出力ニューロンは互いに抑制性の相互作用をしていると考える。この抑制性相互作用は、すべて同じ強さであると仮定される。この抑制性相互作用によって、より大きい値をまず出力したニューロンは、他のニュー



ロンの出力をより強く抑制する, というような, 出力ニューロン間での競合が起こる. その結果, はじめに最も大きい値を出力していたニューロンだけが最終的に大きい値を出力し, その他のニューロンはすべて, ほとんど0に等しいような値を出力するような状態へと, やがて落ちつくようになる. このようにして, 手持ちの参照ベクトルが入力信号にもっとも近いようなニューロンが,  $N$  個のニューロンのなかから選び出される.

そうしてから, Hebb 則にしたがう形での, 参照ベクトルの更新がおこなわれる. 競合によって, 手持ちの参照ベクトルが入力信号にもっとも近いようなニューロンだけが大きい値を出力しているから, このニューロンだけが Hebb 則によって手持ちの参照ベクトルを更新すると考えることができる. これが, 式 (4.1) の更新規則に対応している.

次に, Kohonen のモデルについて考える. Kohonen のモデルでも, 抑制性相互作用は存在すると考える. それに加えて, 近傍に属するニューロンには, 興奮性の相互作用ももっていると考えるのである. これらの相互作用によって, はじめに大きい値を出力していたニューロンとそのニューロンの近傍に属するニューロンとが協調して最終的に大きい値を出力し, 他のニューロンはほとんど0に等しい値を出力するような状態に落ちつくものとする.

そうしてから, Hebb 則にしたがう形での, 参照ベクトルの更新がおこなわれる. 競合によって, 手持ちの参照ベクトルが入力信号にもっとも近いようなニューロンが大きい値を出力しているが, それに加えて, 協調によって, その近傍のニューロンも大きい値を出力しているから, これらのニューロンだけが Hebb 則によって手持ちの参照ベクトルを更新すると考えることができる. これが, 式 (4.3) の更新規則に対応している.

ここで述べたような, 抑制性相互作用による競合, および近傍内での興奮性相互作用による協調のダイナミクスは, 相互作用の大きさがすべて対称である場合には, Hopfield [33, 34] が提案した相互結合形の神経回路網のダイナミクスと本質的に同じである. したがって, 十分時間が経過すれば, 系の状態は安定平衡状態に落ちつくことは, 系のダイナミクスがポテンシャル関数をもつことによって, 理論的に保証されている. 相互作用の影響が十分大きくなるようにすれば, はじめに最も大きい値を出力していたニューロン, あるいはその近傍に属するニューロンだけが最終的に大きい値を出力するような状態に落ちつくようになることが期待できる.

実際の神経系で、このような相互作用が存在するかどうか、ということは、検討されなければならない重要な問題である。神経生理学などでの知見によると、現象としてはこのような相互作用に類似したものは観察されるけれども、抑制性相互作用に直接に対応するようなニューロンの接続様式はみつかっていない。したがって、ここで述べたような抑制性相互作用は、直接的な抑制性相互結合によらないで説明されなければならない。Reggiaら [72] は、この問題を取り扱っているが、彼らの議論によれば、抑制性相互作用の広範囲の形が直接的な抑制性相互作用によらない形へ翻訳できるので、抑制性相互作用が実際にどのような機構で実現されているかは、系の動作を知るうえでは現時点ではそれほど重要ではない。したがって、本論文ではこの問題についてこれ以上は議論しない。

### 4.3 平均学習方程式

式 (4.1) および式 (4.3) で表される学習方程式を解析すれば、これらのモデルの性質がわかることになる。けれども、これらの式はこのままでは取り扱いが困難である。ここで、理論的な取り扱いを容易にするために、これらの学習方程式を統計的に平均化することで得られる平均学習方程式を導入する。

式 (4.1) において、 $\varepsilon$  が十分小さいとしよう。すると、式 (4.1) の両辺に対して、統計的平均をとるという操作をすることができると考えられる。その結果として、時間軸のスケールを適当に取り直して、以下のような式が得られる。

$$\dot{w}_i = \int_{\mathbf{x} \in V_i} \{D(\mathbf{x} - \mathbf{w}_i) e_{i(\mathbf{x})}\} p(\mathbf{x}) d\mathbf{x} \quad (4.9)$$

ここで、 $V_i$  は「 $w_i$  が最近接の参照ベクトルとなるような入力信号  $\mathbf{x}$  の領域」である。この式を、基本的な競合学習に対する平均学習方程式という。

参照ベクトル  $w_i$  は、短期的にみればいろいろな方向に動き回るが、長い時間スケールでならしてみれば、「領域  $V_i$  の、 $p$  による重みつき重心」に向かって移動すると考えてよい (図 4.4)。平均学習方程式は、直観的にはそのような考えを表現している。

Kohonen のモデルに関しても、同様な議論ができる。εが十分小さいとすると、式 (4.3) から次の平均学習方程式が得られる。

$$\dot{w}_i = \int_{C_i} \{D(x - w_i)e_i\} p(x) dx \quad (4.10)$$

ここで  $C_i$  は、 $C_i = \bigcup_{j \in N_j} V_j$  で定義される領域である。いいかえれば、 $C_i$  は「入力信号  $x$  が  $C_i$  の内部に発生したときにニューロン  $i$  が学習する」ような領域である。式 (4.10) を、Kohonen のモデルに対する平均学習方程式という。

基本的な競合学習に対する平均学習方程式 (4.9) は、次の式で表されるポテンシャル関数をもつことが知られている [38].

$$E = \int L(x - w_{i(x)}) p(x) dx \quad (4.11)$$

ここで  $L$  は、 $|\nabla L(z)| = D(z)$  をみたす関数である。通常用いられる距離の尺度として  $(r-1)$  乗ノルム  $D_r(z) = \|z\|^{r-1}$  によるものを考えるならば、 $L_r(z) = \|z\|^r / r$  である ( $r > 0$ )。基本的な競合学習は、式 (4.11) のポテンシャル関数  $E$  に関して確率的最急降下法を行っていることになる。

$(r-1)$  乗ノルム  $D_r(z) = \|z\|^{r-1}$  による距離の尺度をとったとき、ニューロン数  $N$  が十分大きいとき、 $S$  でみたときの参照ベクトルの漸近的な分布  $q(x)$  は変分法によって求めることができ、

$$q(x) \propto p(x)^{n/(n+r)} \quad (4.12)$$

となることが分かっている [50]。さらに第 2 変分を実際に計算することにより、この解が  $r > 0$  で少なくとも局所的に安定であることを容易に確かめることができる。すなわち、 $r$  乗誤差を最小とするためには  $(r-1)$  乗ノルムを使って基本的な競合学習を行えばよく、学習の結果として参照ベクトルの分布は入力信号の分布の  $n/(n+r)$  乗を近似する。この結果は  $N$  が十分大きい場合で考えるならば  $\|z\| \rightarrow 0$  で  $D(z) \approx \|z\|^{r-1}$

とみなせるような一般のノルム  $D(z)$  による距離の尺度をとった場合にも漸近的に成立すると考えてよい。

当然ながら、式 (4.12) の結果は、最適なベクトル量子化器がみだすべき統計的条件として、Zador [97] によって求められた結果と全く同一のものである。

Kohonen のモデルでは、学習が進むにつれて以下の2つの過程が進行するとされている。

(a) 参照ベクトルの  $S$  での配置が  $F$  でのニューロンの配列に従うように、 $S$  での参照ベクトルの整列がおこる。

(b) 参照ベクトルの配置が学習の平衡状態にむかって収束していく。

(a) の性質は位相保存写像の自己組織的形成という観点から重要であるが、整列が確実に起こることは1次元入力の場合に理論的に保証されるに過ぎず、一般の場合には実験的に確認されているにとどまる。本論文では、参照ベクトルの整列については1次元の場合についてあとで議論することにし、当面の議論は、参照ベクトルの整列はすでに完了しているとして進める。

Kohonen のモデルによって形成される参照ベクトルの分布を考察するためには、式 (4.10) によって定まる力学系のダイナミクスについて解析すればよい。基本的な競合学習に対しては式 (4.11) の形のポテンシャル関数が存在したので、それを使うことによって形成される参照ベクトルの分布を考察できた。しかし、Kohonen のモデルに対してポテンシャル関数類似の関数を使って性質を考察することはなされてはいる [93] もの、式 (4.11) に正確に対応するポテンシャル関数は見いだされていない。

本論文では、平均学習方程式 (4.10) の平衡条件  $\dot{w}_i = 0$  を使って、形成される参照ベクトルの漸近的分布を導く。

## 4.4 参照ベクトルの漸近的分布

### 4.4.1 モデルが1次元の場合

モデルが1次元、つまり  $n = 1$  の場合、入力信号は1次元であり、 $F$  におけるニューロンの配列としても1次元のものを考えている。 $F$  のなかでのニューロン  $i$  の近傍  $N_i$

として

$$N_i = \{j \mid |i - j| \leq m\} \quad (4.13)$$

をとると、近傍の大きさ  $m$  によって参照ベクトルの分布がきまる。

具体的には、距離の尺度としてノルム  $D_2$  によるものを取り、

$$q(x) \propto p(x)^s \quad (4.14)$$

の形にあらわしたときに、

$$s = \frac{1}{3} \left[ 2 - \frac{1}{m^2 + (m+1)^2} \right] \quad (4.15)$$

となることが、理論的検討および計算機シミュレーションによって確認されている [75, 90]. また、距離の尺度として一般にノルム  $D_r$  によるものをとったときには、文献 [75] において Ritter が式 (4.15) の結果を導いたのと全く同様にして、簡単に次の結果を得ることができる。

$$s = \frac{1}{1+r} \left[ 2 - \frac{1}{m^2 + (m+1)^2} \right] \quad (4.16)$$

式 (4.16) の形をみると、まず  $m = 0$  のときは  $s = 1/(1+r)$  であるが、 $m$  が大きくなるにつれて  $s$  の値は速やかに  $2/(1+r)$  へ近づき、 $m \geq 1$  では  $m$  の値によらずほぼ  $s = 2/(1+r)$  であると考えてよい。このことは、参照ベクトルの漸近的分布の形が、近傍学習を「するかしないか」によって大きく影響を受けるが、近傍の大きさ  $m$  にはあまり依存しない、ということを示している。

#### 4.4.2 モデルが $n$ 次元の場合

前節と違ってモデルの次元が一般に  $n$  である場合には、近傍の大きさを明示的に取り扱うことはきわめて困難である。けれども、前節での 1 次元での結果からの類推によって、 $n$  次元の場合でも、近傍学習を「するかしないか」によって結果が大きく違ってく

るが、近傍学習をする場合には、近傍の具体的な大きさによって結果が大きく違ってく  
 るようなことはない、と想定される。近傍学習をしない場合についての結果は、4.3で  
 述べたようにすでに得られている。そこで本節では、近傍はかなり大きくとっていると  
 したときの近似的結果を導く。

$F$ でのニューロンの配列としては、 $n$ 次元立方格子状の配列を考える。ひとつのニュー  
 ロンは、 $n$ 個の添字の組  $i = (i_1, i_2, \dots, i_n)^T$  で指定される。ニューロン  $i$  の近傍  $N_i$   
 として、次のような(超)立方体状のものを考える。

$$N_i = \{i + j \mid j = (j_1, j_2, \dots, j_n)^T, \max_{1 \leq k \leq n} |j_k| \leq m\} \quad (4.17)$$

学習が十分進んだ状態を考え、参照ベクトルは整列を完了しているとする。入力確率密  
 度分布  $p(x)$  が十分なめらかであれば、参照ベクトルの密度分布も十分なめらかである  
 とみなしてよいであろう。いま、ある参照ベクトル  $w_i$  に注目する。 $x = w_i$  の付近で  
 は、ニューロン  $i$  の近傍  $N_i$  に含まれるニューロンの参照ベクトルは  $F$  でのニューロン  
 の配列の様子を保って配列している。 $x = w_i$  において、参照ベクトルの密度は  $q(w_i)$ 、  
 密度勾配は  $\nabla q(w_i)$  である。これを次のようにモデル化する。

まず0次近似として、密度勾配  $\nabla q(w_i)$  は考えないでモデル化を行う。ニューロン  $i$   
 の近傍  $N_i$  に含まれる  $(2m+1)^n$  個のニューロンの参照ベクトルは、 $S$  のなかで  $x = w_i$   
 を中心として(超)立方体状に配列していると考えられる。最近接の参照ベクトル間の  
 距離を  $d$  とすると、 $d^{-n} = q(w_i)$  である。 $x = w_i$  を原点として適当な正規直交基底  
 $\{e_1, e_2, \dots, e_n\}$  をとり、ニューロン  $(i+j) \in N_i$  の参照ベクトルの0次近似での位  
 置  $w_{i+j}^{(0)}$  を、 $w_i$  からみた相対的な位置関係  $\Delta w_j^{(0)} = w_{i+j}^{(0)} - w_i$  であらわすと、

$$\Delta w_j^{(0)} = d(j_1 e_1 + j_2 e_2 + \dots + j_n e_n) = dUj \quad (4.18)$$

$$U = [e_1, e_2, \dots, e_n] \quad (4.19)$$

となる。

次に密度勾配  $\nabla q(w_i)$  を考慮した1次近似を考える. 式(4.18)であらわされる(超)立方体状の配列が, 密度勾配の存在によって摂動をうけると考えてモデル化する.  $\nabla q(w_i)$  の向きの単位ベクトルを  $e_g$  とする. 0次近似の  $\Delta w_j^{(0)}$  は, その  $e_g$  方向の成分の大きさ  $e_g \cdot \Delta w_j^{(0)}$  に比例して次式にしたがって1次近似としての値  $\Delta w_j$  へ修正されると考える.

$$\Delta w_j = (1 - ce_g \cdot \Delta w_j^{(0)}) \Delta w_j^{(0)} \quad (4.20)$$

式(4.20)による参照ベクトル配置の修正例を図4.5に示す.

比例係数  $c$  の値は, 式(4.20)に従っている配列の密度勾配が  $\nabla q(w_i)$  と一致するという条件から定めることができ, 次式で与えられる.

$$ce_g = \frac{1}{n+1} \frac{\nabla q(w_i)}{q(w_i)} \quad (4.21)$$

距離の尺度として  $(r-1)$  乗ノルム  $D_r$  によるものを使うことにしよう. 平均学習方程式による, 参照ベクトル  $w_i$  に関するつりあいの条件  $\dot{w}_i = 0$  は, 近傍の大きさ  $m$  が十分大きいときには近似的に次式で与えられる.

$$\int_C |x - w_i|^{r-2} (x - w_i) p(x) dx = 0 \quad (4.22)$$

ここで,

$$C = g(C') \quad (4.23)$$

$$g(z) = w_i + f(z) \quad (4.24)$$

$$f(z) = (1 - ce_g \cdot z)z \quad (4.25)$$

$$C' = [-(m + \frac{1}{2})d, (m + \frac{1}{2})d]^n \quad (4.26)$$

である.

領域  $C$  の境界は一般には入り組んだ形をしているが、参照ベクトルが局所的に整列していれば、領域  $C$  をなめらかな境界をもつ領域  $g(C')$  によって近似することができると考えてよい (図 4.6). ここでおこなっているのは、そのような近似である.

変数変換  $x = g(z)$  によって変数を  $z$  にかえると,

$$\int_{C'} |f(z)|^{r-2} f(z) p(\mathbf{w}_i + f(z)) J[f] dz = 0 \quad (4.27)$$

となる. ここで,  $J[f]$  は  $f$  のヤコビアンである.  $p(x)$ ,  $f(z)$  がともに十分緩やかであるとして, 式 (4.27) の左辺の被積分関数を  $z = 0$  のまわりにテイラー展開し,  $z$  について高次の項を無視するという近似のもとで計算すると,

$$K [\nabla p(\mathbf{w}_i) - c(n+r)p(\mathbf{w}_i)e_g] = 0 \quad (4.28)$$

という結果が得られる. ここで,  $K$  は学習に使う近傍の大きさと信号空間の次元  $n$  とだけに依存する定数である. 式 (4.28) の詳しい導出は付録 A に示しておく.

この結果と式 (4.21) とから, 平均学習方程式のつりあいの条件は

$$\frac{\nabla q(\mathbf{w}_i)}{q(\mathbf{w}_i)} = \frac{n+1}{n+r} \frac{\nabla p(\mathbf{w}_i)}{p(\mathbf{w}_i)} \quad (4.29)$$

である. これは  $x = \mathbf{w}_i$  のところで成立すべき条件式である. けれどもニューロン数  $N$  が十分大きい場合を考えるならば, この条件式は入力信号空間  $S$  のいたるところで成立していると考えてよい. こうして以下の微分方程式を得る.

$$\frac{\nabla q}{q} = \frac{n+1}{n+r} \frac{\nabla p}{p} \quad (4.30)$$



これを解けば, Kohonen のモデルが形成する参照ベクトル分布  $q$  と入力信号の確率密度分布  $p$  とのあいだの定量的関係をあらわす結果が得られる [92].

$$q(\boldsymbol{x}) \propto p(\boldsymbol{x})^{(n+1)/(n+r)} \quad (4.31)$$

ここで得られた結果は,  $n = 1$  のときの結果 (式 (4.16)) で  $m \rightarrow \infty$  とした場合を含み, その拡張になっている. また, この結果は密度勾配を考慮した参照ベクトルの配置の具体的なモデル化のしかた (式 (4.20) あるいは式 (4.25) の  $f(z)$ ) に依存しているように見えるがそうではなく, 付録 A における導出にみるように本質的には  $f(z)$  の  $z$  に関する展開における低次項の形と, 近傍の形の対称性とだけから式 (4.31) と同じ結果が得られるので, この結果自体はより一般的な条件のもとで成立するものと考えてよい.

ここでは, 局所的なつりあいの条件から参照ベクトルの分布  $q(\boldsymbol{x})$  の形を導いた. けれども  $n \geq 2$  の場合には, 参照ベクトルの配列の大域的な様子によって局所的な参照ベクトルの配列が影響を受けることが考えられる. 本論文ではそのような効果は考慮していないことを注意しておく.

#### 4.4.3 計算機シミュレーション

この節では, 前章での理論的結果を確認するために計算機シミュレーションを行う. モデルが 2 次元, すなわち入力信号が 2 次元でなおかつ参照ベクトルが 2 次元格子状の配列をもっている場合を対象とする.

入力信号  $\boldsymbol{x} = (x_1, x_2) \in [0, 1]^2$  は確率密度分布  $p(x_1, x_2) = 3x_1x_2^{1/2}$  に従って独立に生起するとする. 参照ベクトルは 2304 個とし,  $48 \times 48$  の配列をなすとする. 参照ベクトルの初期配置は  $q(\boldsymbol{x}) \propto p(\boldsymbol{x})$  となるように, いいかえれば初期状態として最大エントロピー符号化が実現されているようにしておく. 距離の尺度としては,  $r = 2$  すなわち  $D_2(z) = \|z\|$  によるものとする.

学習の収束を早めるために, 学習パラメータ (式 (4.1) の  $\varepsilon$ ) と近傍の大きさを, 学習が進むにつれて小さくしていくようにする. ここでは近傍学習の効果をみるために,

近傍の大きさは最終的に  $m = 3$  となるようにしている。学習の各段階で、 $x_1, x_2$  にかんする参照ベクトルの周辺分布を求め、端の効果を除くために中央の  $1/3$  の部分をとって、回帰分析により  $q(x) \propto \{p(x)\}^s$  と表したときの指数  $s$  の値を求める。

以上の条件で計算機シミュレーションを行った結果を図 4.7 に示す。図 4.7 では、学習の各段階での参照ベクトルの配置を示している。端の効果はみられるものの、参照ベクトルの配列の中央部では安定な配置が形成されている様子がみられる。図 4.8 は、学習の各段階で回帰分析によって求めた指数  $s$  の値を、学習回数に対して表したものである。ここでは  $n = 2, r = 2$  だから、基本的な競合学習であれば  $s = 1/2$ 、近傍学習を考慮した前章での結果によれば  $s = 3/4$  であり、近傍学習を行っているこの例では学習が進むにつれて指数  $s$  の値は  $3/4$  に近づいていくことが予想される。 $x_1, x_2$  それぞれに対する結果をみると、お互いに多少の食い違いはあるものの、どちらも学習が進むにつれて  $s = 3/4$  に近づいていく傾向がみられる。以上から、前章で得られた理論的結果の妥当性が確認できたと考えられる。

## 4.5 学習の収束性

### 4.5.1 概要

Kohonen のモデルに関する問題点として、学習がいつも成功するという保証がないことが挙げられる。本論文でも、前節までに、一般の次元に対して、Kohonen のモデルが形成する参照ベクトルの分布と入力信号の確率密度分布との関係をあらわす結果を近似的に得た。ただ、前節での結果も、得られた結果を計算機シミュレーションによって確認しているに過ぎない。一般的に言って、Kohonen のモデルに関する理論的な研究では、学習をおこなうことによって、モデルの状態が理論的に導かれた平衡状態へ安定に収束するかどうか、ということに対する考察があまりなされていない。

このような問題点に関連して、Kohonen のモデルが実際に学習に失敗した計算機シミュレーション例がいくつか示されている [52, 62, 96]。学習の失敗には、大域的な構造の形成に失敗する場合と、微視的に一様な構造が不安定化する場合とがある。理論的には、大域的な構造の形成は、次節で議論するように、一次元の場合に保証されるに過ぎないが、一般の場合にも近傍学習に関するアルゴリズムを工夫すれば実際的には大域的

な構造の形成を促進できる, という結果 [62] などが得られている. 一様な構造の不安定化については, 一次元一様入力の場合に近傍学習に関するアルゴリズムの設定次第で不安定化が生じる, という結果 [52, 96] がある. どちらの場合に対しても現状では, 倉田のいう「なで肩」の重み, あるいは Lo らのいう neighborhood function (どちらも同じことを指していると考えてよい) を使うことによって学習の失敗を減少させることができると考えられている.

本節では, 上記の意味での学習の失敗がないとした場合に, Kohonen のモデルが形成する参照ベクトルの分布が式 (4.31) で示したような分布に安定に収束することを示す. そのために, Kohonen のモデルに関する平均学習方程式に対して, 粗視化した意味での Lyapunov 関数を与えることで, Kohonen のモデルが粗視化してみたときには安定に収束することを示す [91].

#### 4.5.2 学習方程式の流体近似

本節では, 式 (4.28) と同様にして求めた平均学習方程式の近似式をさらに流体近似によって流体力学的方程式にかきなおす. それに基づいて Kohonen のモデルが式 (4.31) で示したような結果へ安定して収束することを示す.

平均学習方程式の線形近似は, 位置  $\mathbf{x}$  における参照ベクトルに対しては, 式 (4.28) と同様にして次のような形にあらわされる.

$$\dot{\mathbf{x}} = K \left\{ \nabla p(\mathbf{x}) - \frac{1}{s} \frac{\nabla q(\mathbf{x})}{q(\mathbf{x})} p(\mathbf{x}) \right\} \quad (4.32)$$

次に, 参照ベクトル分布  $q(\mathbf{x})$  自体の時間変化について考える. 参照ベクトルの数は学習を通じて一定だから,  $q(\mathbf{x})$  を信号空間全体にわたって積分したものは, 時刻にかかわらず一定値である. 本稿では, この一定値が 1 になるように  $q$  が規格化してあるとする. すると,  $q$  は質量を保存する流体の密度をあらわしているというようにも考えることができる. 流体力学における連続の方程式は, 次の形をしている.

$$\dot{q} + q \nabla \cdot \dot{\mathbf{x}} = 0 \quad (4.33)$$

式 (4.33) によって、各点  $x$  での参照ベクトルの移動方向  $\dot{x}$  と、参照ベクトル分布の時間変化  $\dot{q}$  とが関係づけられている。したがって、式 (4.32) と式 (4.33) とを併せて考えると、 $q$  の時間変化をあらわす閉じた表式が得られたことになる。

### 4.5.3 Lyapunov 関数

本節では、4.5.2 節で説明した、参照ベクトルの分布の時間変化  $\dot{q}$  に対する Lyapunov 関数を与える。

$c$  を、 $cp^s$  が信号空間全体で積分して 1 になるような規格化定数とする。 $q$  は規格化されていると仮定している。4.4 節での結果は、時間的に変化する確率密度分布  $q$  の平衡分布が確率密度分布  $cp^s$  であることを示している。 $q$  が  $cp^s$  へ収束することを示すためには、これら 2 つの確率密度分布のあいだの距離のようなものを考えると都合がよい。本節では、2 つの確率密度分布  $p_1$ ,  $p_2$  のあいだの「距離」 $D$  が以下の式であらわされるとする。

$$D(p_1, p_2) = \int_S f\left(\frac{p_1}{p_2}\right) dx \quad (4.34)$$

ここで  $f(z)$  は  $z > 0$  に対して定義された関数で、

$$\begin{cases} f(1) = 0 \\ z \neq 1 \longleftrightarrow f(z) \neq 0 \\ f'(1) = 0 \\ (zf'(z))' > 0 \end{cases} \quad (4.35)$$

をみたすものとする (例えば、 $f(z) = (\log z)^2$  などがこの条件をみたす)。すると  $D$  は、 $p_1 \equiv p_2$  のとき、かつそのときに限り 0 になり、それ以外のときは  $D > 0$  である。

Kohonen のモデルに対して、汎関数  $J[q]$  を次のように定義する。

$$\begin{aligned} J[q] &= D(q, cp^s) \\ &= \int_S f\left(\frac{q}{cp^s}\right) dx \end{aligned} \quad (4.36)$$

$q/(cp^s) = X$  とおく. この被積分関数の  $q$  に関する変分導関数  $\delta J$  は以下のようになる.

$$\begin{aligned}\delta J &= \frac{\partial}{\partial q} \left( f \left( \frac{q}{cp^s} \right) \right) \\ &= -\frac{1}{q} X f'(X)\end{aligned}\quad (4.37)$$

Kohonen のモデルの学習にしたがって  $J$  がどのような時間変化をするかを調べるために,  $J$  の時間微分  $\dot{J}$  を計算する. 付録 B に示した導出によって,

$$\begin{aligned}\dot{J} &= \int_S \delta J \cdot \dot{q} \, dx \\ &= -\frac{K}{s} \int_S [X f''(X) + f'(X)] \nabla X \cdot \left( p \frac{\nabla X}{X} \right) \, dx \\ &= -\frac{K}{s} \int_S \left[ f''(X) + \frac{f'(X)}{X} \right] p \|\nabla X\|^2 \, dx\end{aligned}\quad (4.38)$$

という結果を得る. 式 (4.35) に注意すると, 式 (4.38) から  $\dot{J} \leq 0$  であることがいえる. また,  $\dot{J} = 0$  となるのは,  $\nabla X = 0$  すなわち  $q = cp^s$  のときに限ることもわかる.

さて,  $q = cp^s$  のときは  $J = 0$  であり, それ以外のときは  $\dot{J} < 0$  であることがわかったのだが, このことは,  $J$  が Kohonen のモデルに対する Lyapunov 関数になっていることを示している. 従って, Kohonen のモデルにおける学習によって参照ベクトル分布は  $q = cp^s$  という分布に収束することが示されたことになる.

#### 4.6 参照ベクトルの整列に関する検討

Kohonen のモデルでの参照ベクトルの整列は, Kohonen のモデルの性質のなかでも重要である. 参照ベクトルの整列は, Kohonen のモデルを生体の神経系に関するモデルとしてみる場合には, 大域的な位相保存写像の形成に関連する. また, 工学的な適応形信号処理アルゴリズムとしてみる場合には, 信号源の大域的な構造をとりだすことに関係する.

計算機シミュレーションによれば, Kohonen のモデルでの参照ベクトルの整列に関しては, もともとのアルゴリズムよりも, 参照ベクトルを更新する際に参照ベクトルごとに適当な重みをつけて更新をおこなうような方法のほうが, 概してよい結果を示す.

Kohonen のモデルでは, 入力信号  $x$  が発生したときニューロン  $i(x)$  の近傍  $N_{i(x)}$  に含まれるニューロン  $j$  が学習するのだが, いま, 大きさ  $m$  の近傍を  $N_{i(x)}^{(m)}$  であらわし,  $j \in N_{i(x)}^{(m)}$  でありかつ  $j \in N_{i(x)}^{(m-1)}$  であるとしよう. このようなニューロンを学習させるときには,  $m$  の関数として与えられる重み  $h(m)$  にしたがって参照ベクトルの更新をおこなうような方法を考える.

$$\Delta w_j = \varepsilon h(m) D(x - w_j) e_j \quad (j \in N_{i(x)}^{(m)} \text{ かつ } j \notin N_{i(x)}^{(m-1)}) \quad (4.39)$$

このような, 近傍にしたがって重みをつける形の学習則を, 重みつき学習則ということにする. また, 関数  $h(m)$  を近傍関数ということにする. 重みつき学習則からみると, これまで議論してきた, 重みをつけない形の学習則は, 学習に用いる近傍の大きさを  $m_0$  とすると  $m \leq m_0$  で  $h(m) = 1$  であり  $m > m_0$  では  $h(m) = 0$  であるような学習則として, すなわち重みつき学習則の特別な場合として捉えることができる.

重みつき学習則は, モデルが 1 次元の場合については理論的にもよく調べられている. 例えば, 近傍関数  $h(m)$  が  $m$  に関して単調減少であれば, 入力信号が一様な確率密度分布にしたがって生起する場合には, 参照ベクトルの平衡配置が安定になることがわかっている [52, 96]. また, 近傍関数  $h(m)$  が  $m$  に関して狭義単調減少であれば, 確率 1 で参照ベクトルが整列することが証明されている [11, 16, 17]. 単調減少の近傍関数は, これらの理論的結果が得られる以前から, 経験的に有効な方法として利用されてきている [63, 73, 74]. モデルの次元が 2 以上の一般の場合に関して重みつき学習則の性質を理論的に検討しようという試みもなされている [62] が, 満足な成果は得られていないのが現状である [17].

この問題に関して本論文では, Erwin らの議論 [17] を拡張した以下の結果を与える.

定理 1 次元の重みつき学習において、入力信号が有限区間  $[a, b]$  上に発生し、入力信号の確率密度分布がこの区間で 0 になることはないとする。また、参照ベクトル数を  $N$  とし、近傍関数  $h(m)$  に関して、 $1 \geq h(0) \geq h(1) \geq h(2) \geq \dots \geq h(N-1)$  かつ、少なくともどこか 1 ヶ所で等号が不成立、 $h(m) \geq 0$ 、 $h(1) > 0$  という条件が成り立つとする。このとき、参照ベクトルの任意の初期配置に対して、有限の長さの入力系列からなる有限の確率測度をもつ集合で、この集合に属するどの入力系列によっても参照ベクトルが整列するようなものが存在する。

一般に Kohonen のモデルのアルゴリズムは、モデルの現在の状態が与えられれば次にモデルがどのような状態をとるかは確率的に決定されるから、単純マルコフ過程であると考えることができる。したがって、上の定理とアルゴリズムのマルコフ性とから、確率 1 で参照ベクトルが整列することが導かれる。定理の証明は、付録 C に示す。

## 4.7 考察

この章では、初期視覚系での位相保存写像や刺激の頻度に対応した写像の拡大率といった、構造の大域的な特徴が自己組織的に形成されることに関するモデルとして、Kohonen のモデルをとりあげ、その性質について理論的に考察した。刺激の頻度に対応した写像の拡大率については、Kohonen のモデルでは、入力信号の確率密度分布とモデルが形成する参照ベクトル分布との関係という形で取り扱うことになる。この問題に対して、本論文では Kohonen のモデルに対する平均学習方程式を使って議論した。

式 (4.1) や式 (4.3) などの学習方程式からそれぞれに対する平均学習方程式を導くときに、天下りの係数  $\varepsilon$  の値が十分小さければよいと述べたが、この点に関してはもう少し詳しく議論する必要がある。

この問題は、より一般的な立場から、確率微分方程式の平均化操作の妥当性の問題として表現される。具体的にいうと、確率微分方程式を平均化操作することによって導かれる決定論的な微分方程式の解が、もとの確率微分方程式の解とどのように関係するのか、という問題として表現される。Geman [24] はこの種の確率微分方程式の平均化操作に関して議論している。その結果を本論文での議論に当てはめると、本論文で説明したような Kohonen のモデルに対しては、 $\varepsilon$  が十分小さいという条件があれば、平均学

習方程式の解はもとの確率微分方程式の解のよい近似になっていることが保証される。とくに、平均学習方程式の漸近安定な状態が、もとの確率微分方程式の漸近安定性に対応することがいえる。このことによって、 $\varepsilon$  が十分小さいとしたときの Kohonen のモデルの性質を平均学習方程式によって解析するという方法が正当化される。

平均学習方程式は、もとの確率微分方程式の決定論的な近似である。Kohonen のモデルのもともとの確率的な側面を少しでも残して近似するとすれば、拡散近似が得られる。アルゴリズムの確率論的な性質を取りいれて議論するために、拡散近似を使って解析をおこなうほうがよいのではないか、ということも考えられる。けれども拡散近似による議論は、平均学習方程式による議論と比べて格段に困難である。これについて述べる。

拡散近似では、学習の過程は拡散過程で近似される。学習方程式は、以下のような形で表される。

$$dw_i = \varepsilon K_i(\mathbf{w}) dt + \varepsilon^2 (D_i(\mathbf{w}))^{1/2} dB^n \quad (4.40)$$

ここで、 $B^n$  は  $n$  次元ブラウン運動、 $K_i \in \mathbb{R}^n$ 、 $D_i(\mathbf{w}) \in \mathbb{R}^n \times \mathbb{R}^n$  で、

$$\begin{aligned} K_i(\mathbf{w}) &= \int_{C_i} (\mathbf{x} - \mathbf{w}_i) p(\mathbf{x}) d\mathbf{x} \equiv (\mathbf{c}_i - \mathbf{w}_i) \int_{C_i} p(\mathbf{x}) d\mathbf{x} \\ D_i(\mathbf{w}) &= \int_{C_i} (\mathbf{x} - \mathbf{c}_i) \otimes (\mathbf{x} - \mathbf{c}_i) p(\mathbf{x}) d\mathbf{x} \end{aligned} \quad (4.41)$$

とする。 $\otimes$  はテンソル積である。式 (4.40) で右辺第 2 項を無視すると、平均学習方程式 (4.10) と一致する。

モデルの状態は、 $N$  個の参照ベクトルの組  $\mathbf{w} \equiv (\mathbf{w}_1, \dots, \mathbf{w}_N)$  であらわされる。学習方程式が確率微分方程式であることに対応して、学習の過程はモデルの状態分布  $f(\mathbf{w})$  が時間的に変化する、という形で表現される。拡散近似の学習方程式 (4.40) から、状態分布  $f(\mathbf{w})$  に関する Fokker-Planck 方程式を導くことができる。伊藤の算法を使うことにすると、式 (4.40) に対応する Fokker-Planck 方程式は次のようになる。

$$\dot{f} = -\varepsilon \sum_{i=1}^N \sum_{j=1}^n \frac{\partial}{\partial w_{ij}} [[K_i(\mathbf{w})]_{j1} f] + \frac{1}{2} \varepsilon^2 \sum_{i=1}^N \sum_{j=1}^n \sum_{k=1}^n \frac{\partial^2}{\partial w_{ij} \partial w_{ik}} [[D_i(\mathbf{w})]_{jk} f] \quad (4.42)$$



いま問題にしている確率過程はマルコフ過程だから、式 (4.42) はまちがいでなく定常解をもつ。けれども、有用な解析をおこなうためには、この確率過程について「詳細つりあい」が成立している必要がある。もしそうであれば、定常解をギブス分布の形

$$f \propto e^{-\Phi/\epsilon} \quad (4.43)$$

で求めることができる。Φ は一般化されたポテンシャル関数である。このような形で表すことができれば、模擬焼き鈍し法 [5, 25, 47] と関連づけることで、大域的最適解へ収束させるためのアルゴリズム上の工夫に関する議論が可能である。

けれども、Kohonen のモデルに関しては、「詳細つりあい」は一般には成立しない。モデルの次元が 1 より大きい場合には、「詳細つりあい」の条件は可積分条件と同じなので、求積によって Fokker-Planck 方程式を解くことは事実上不可能である [27]。以上の理由によって、拡散近似により Kohonen のモデルの解析をおこなうことは困難であると考えている。

つぎに、生体の神経系でみられる刺激の頻度に対応した写像の拡大率との関連で、Kohonen のモデルについて得られた性質について考察する。Kohonen のモデルでは、入力信号の確率密度分布  $p$  とモデルが形成する参照ベクトル分布  $q$  との関係という形で議論した。モデルが形成する参照ベクトル分布は  $q \propto p^s$  という形で与えられるという結果が得られた。s の値は正なので、参照ベクトル分布  $q$  は入力信号の確率密度分布  $p$  と正の相関をもつが、このことから Kohonen のモデルは、生体の神経系で刺激の頻度に対応した写像の拡大率が実現されることを定性的には説明できる。けれども、通常使われているアルゴリズムでは指数  $s$  の値は 1 より小さくなる。したがって、Kohonen のモデルでは、生体の神経系に関していわれるような、写像の拡大率に関する比例則を、定量的に説明することはできない。

このことに関しては、以下の 2 つの点を考慮する必要がある。ひとつは、生体の神経系に関していわれている比例則が、どの程度精度のあるものなのかが明確にされていない点である。もうひとつは、Kohonen のモデルは、実際の神経系の構造と比べて大幅な単純化がなされている点である。生体の神経系が示す性質がモデルによって定性的

に再現されることから、モデルを得る際の単純化で、本質的な特徴が保たれていることが期待される。また、定量的な性質がモデルによって再現されないのなら、定量的な性質を議論するためにはモデルを単純化し過ぎているのではないか、ということが考えられる。

Linsker [61] は、理論的に比例則を実現するようなモデルを提案している。Amari の 1 次元神経場モデル [3, 99] も比例則が実現されることを主張しているが、Amari のいう比例則は、本論文でいっているものと内容が若干異なる。Linsker のモデルは、本論文で議論した Kohonen のモデルと同種のモデルということができるが、Kohonen のモデルと比較してより複雑な構成をもっている。また、Linsker のモデルは、条件次第で等角写像を形成できる性質をもつことが理論的に示されている。いっぽうで Kohonen のモデルは、競合学習の時点で入力空間の計量に関する情報が欠落してしまうので、等角写像の形成に関しては不利である。これらを考慮すると、生体の神経系に関するモデルとして考える際には、定量的な再現性の点を問題にするならば、Linsker のモデルのほうが Kohonen のモデルよりも有利である。

それよりもむしろ、本章で得られた結果は、Kohonen のモデルを工学的な問題へ応用する際に重要であると考えている。すでに述べたように、Kohonen のモデルは適応的信号処理システムとして様々な問題に実際に適用されている。Kohonen のモデルがアルゴリズムとして単純であることが、実際の問題への応用という局面では有効にはたらいっている。

適応的ベクトル量子化アルゴリズムとしてみたときには、符号化における  $r$  乗誤差の期待値が小さくなるようなもの、あるいは、得られる符号のエントロピーが大きくなるようなものが実用的には望まれる。これらの性質を参照ベクトル分布  $q$  および入力信号の確率密度分布  $p$  によってあらわすと、 $r$  乗誤差の期待値が最小なベクトル量子化器に関しては  $q \propto p^{n/(n+r)}$ 、また最大の符号のエントロピーを与えるエントロピー符号化器に関しては  $q \propto p$  となる。

Kohonen のモデルに関しては、形成される参照ベクトル分布は  $q \propto p^s$  という形になる、という結果が得られたので、上に述べたことから、 $s = n/(n+r)$  となるアルゴリズムに関しては  $r$  乗誤差の期待値が最小なベクトル量子化器が、また  $s = 1$  となる

アルゴリズムに関してはエントロピー符号化器が、それぞれ得られることになる。簡単のために、Kohonen のモデルにおいて学習に使う近傍の大きさが十分に大きいとしよう。このときには、指数  $s$  の値は  $s = (n+1)/(n+r)$  となる。基本的な競合学習では  $s = n/(n+r)$  である。このことから、基本的な競合学習と同じく Kohonen のモデルでも、入力信号の次元  $n$  が大きくなると得られる符号は漸近的にエントロピー符号に近づく、という性質があることがわかる。さらに Kohonen のモデルで得られる符号のほうがよりエントロピー符号に近い、という性質がある。また、距離尺度を  $r = 1$  となるようにとると、 $n$  の値にかかわらず  $s = 1$  となるので、エントロピー符号を構成しようという目的には、Kohonen のモデルのほうが有利である。

また、粗視化してみたときのアルゴリズムの収束に関して保証が与えられたことにより、上で議論したような、得られる符号の漸近的な性質については、実際の問題にアルゴリズムを適用した場合にも実現されることが期待できる。一方で、Kohonen のモデルの重要な性質のひとつとされている、参照ベクトルの整列に関しては、理論的には依然としてモデルが 1 次元の場合にしか、整列が完了することは保証されていない。この問題は応用上も重要なので、より一層研究がなされなければならないと考えている。

本論文ではとりあげなかったが、入力信号の次元と参照ベクトルの配列の次元とが同じでない場合には、Kohonen のモデルは、特徴次元の自動選択 [49] とよばれる性質を示す。この性質も、ある意味でモデルの安定性にかかわる重要な問題である。Kohonen のモデルを実際の問題に適用することを考えるときには、この性質に関しても検討する必要がある [6, 51, 74].

## 第 4 章 図表

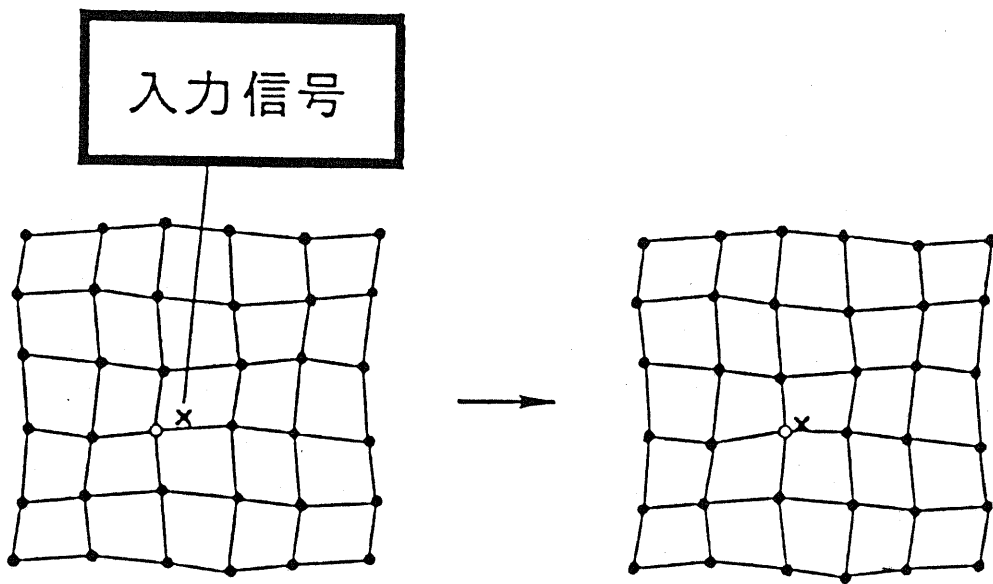
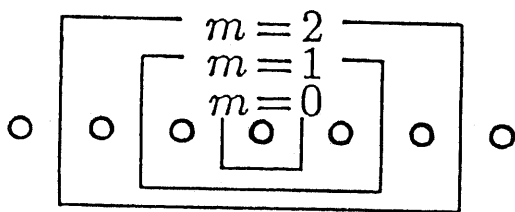
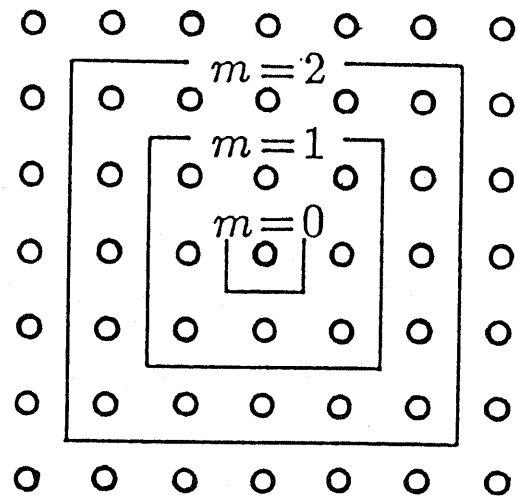


図 4.1 基本的な競合学習



(a) 1次元の場合



(b) 2次元の場合

図 4.2 Kohonen のモデルにおけるニューロンの配列, 近傍の例

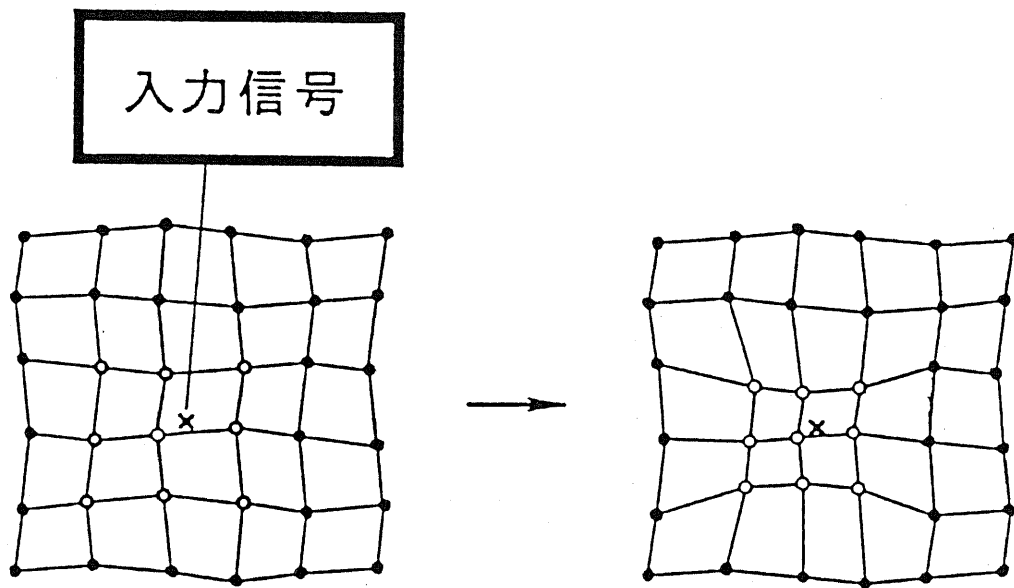


図 4.3 Kohonen のモデルにおける学習

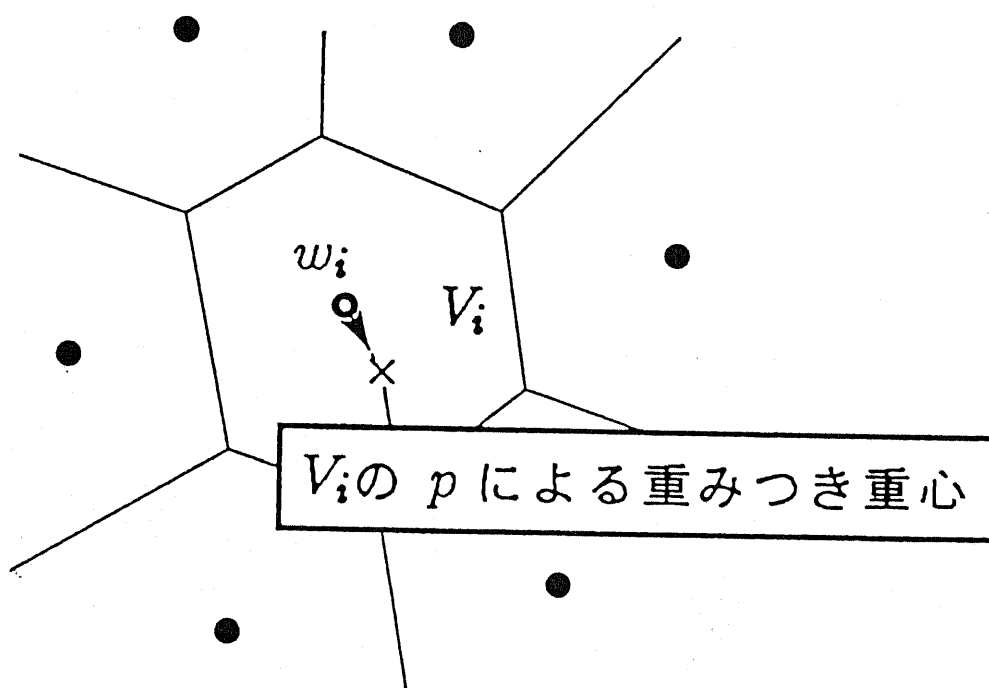


図 4.4 平均学習方程式



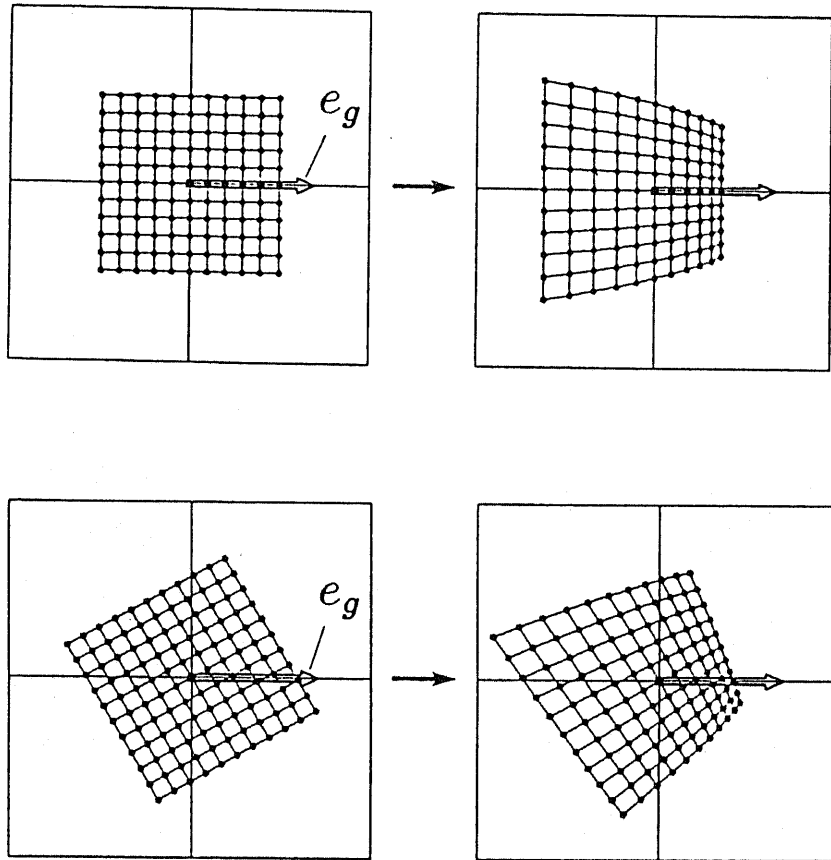


図 4.5 参照ベクトル配置の修正例

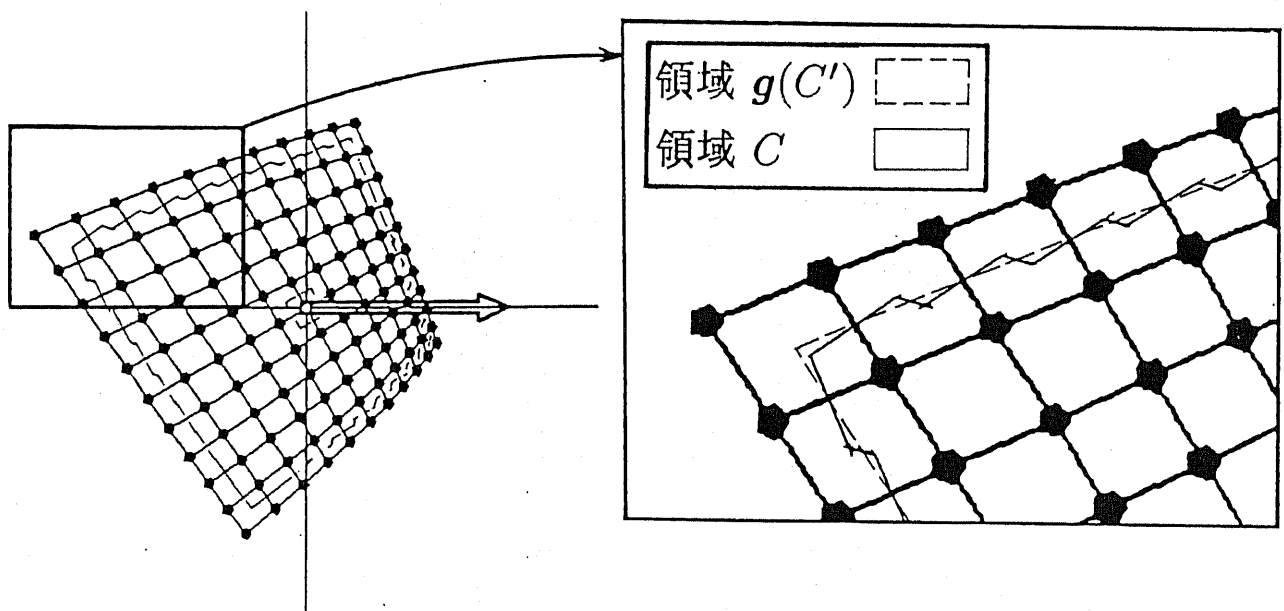
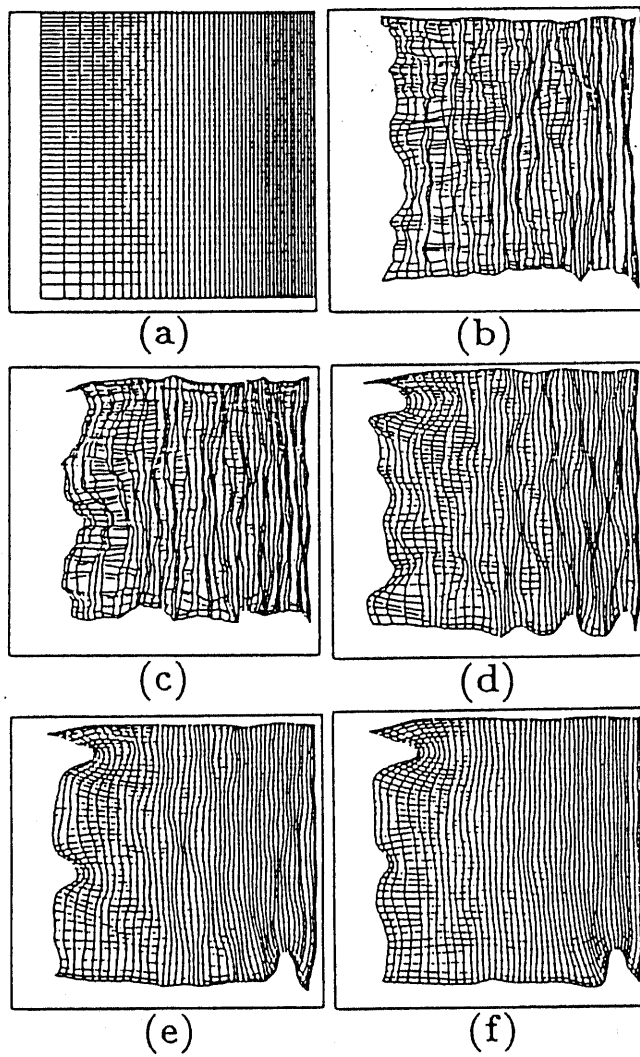


図 4.6 領域  $g(C')$  による領域  $C$  の近似



学習回数 : (a) 0, (b)  $10^3$ , (c)  $10^4$ , (d)  $10^5$ , (e)  $10^6$ , (f)  $5 \cdot 10^6$ .

図 4.7 学習による参照ベクトル配置の変化の様子

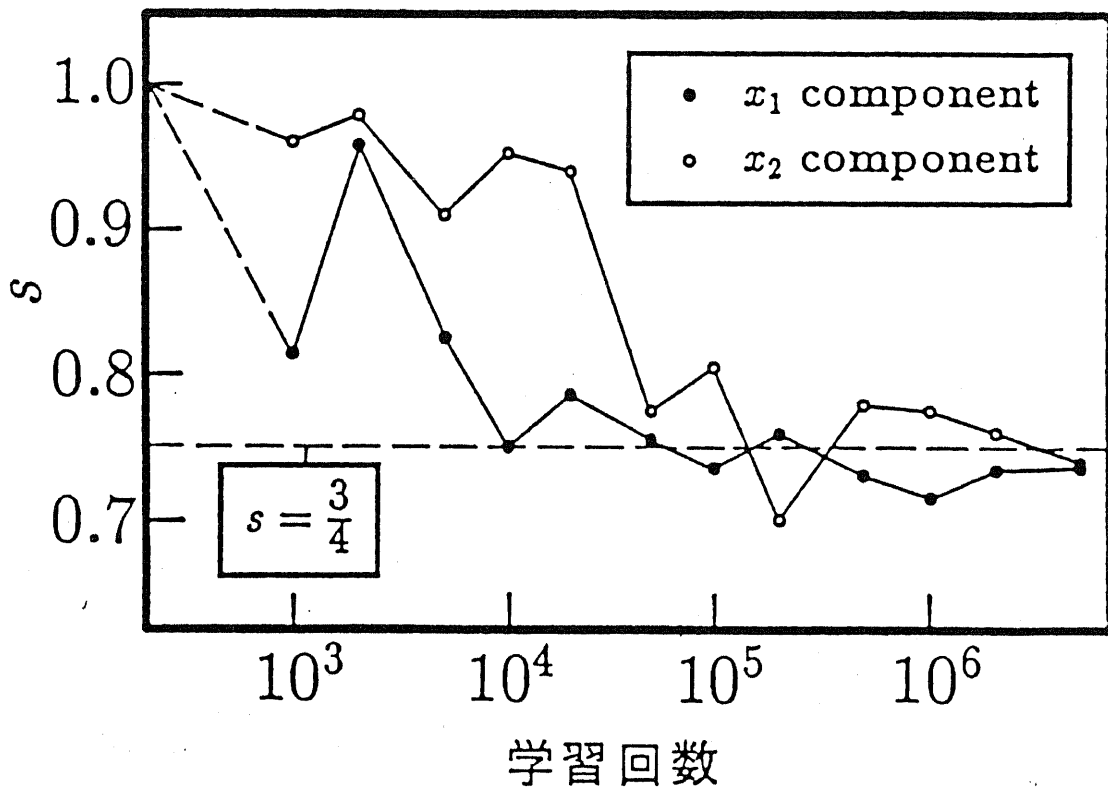


図 4.8 学習による指数  $s$  の変化

## 第 5 章

# 構造の局所的な特徴の自己組織的形成 に関するモデル

心を虚しくして彫りたまえ, 仁王は埋っている  
のだから... ([43])

## 5.1 本章の目的

Linsker [58-60] は、層構造をもつフィードフォワード 2 次元神経場での自己組織化モデルを扱っている。Linsker のモデルでは単純な Hebb 則を学習則として使っているにもかかわらず、1 層目に無相関な雑音入力を与えてモデルを学習させることによって、2 層目には all-on / all-off 形のニューロンができ、3 層目には on-中心形 / off-中心形ニューロンが、そして早ければ 4 層目には方位選択性ニューロンがそれぞれできあがる、というような単純でなく興味深い性質を Linsker のモデルは示す。この章では、Linsker のモデルを、初期視覚系において刺激選択性ニューロンが自己組織的に形成されることに関するモデルとしてとりあげる。第 2 章でみたように、網膜から一次視覚野までは、神経系の構成は基本的に層状フィードフォワードとみなされることが多く、Linsker のモデルはこの立場とは相性がよい。

Linsker [58-60] は彼のモデルで刺激選択性ニューロンが形成されるという性質を単にコンピュータシミュレーションの結果として示し、後で「情報量最大保持原理」としてその意味づけを考察している [57] だけである。MacKay と Miller [64] は、Linsker のモデルの性質を理論的に解析することを試みている。彼らが扱っているのは、3 層目でおこる on-中心形 / off-中心形ニューロンの形成に関してである。本章では、彼らの方向をさらに進めれば方位選択性ニューロンの形成に関しても解析ができるのではないかと考え、それを試みている。

## 5.2 Linsker のモデル

### 5.2.1 基本的な学習の機構

Linsker のモデルは層状、フィードフォワード形で、各層のニューロンは隣接する前層だけから入力をうける。各層で基本的な構成、学習則は同じなので、一層の学習に関する解析結果は、入力の条件が同じなら基本的にはどの層にも適用できる。また、各層は一様な 2 次元神経場をなしていると考えるので、層内の一つのニューロンをとりだし、それと前層とのシナプス結合に関して学習則を適用した場合について解析すれば十分である。

L 層と M 層とを隣接する 2 層とし, M 層のひとつのニューロン  $m$  と L 層の各ニューロンとのシナプス結合に関して考える (図 5.1). 便宜上,  $m$  の位置に対応する L 層上の点を原点にとる.

L 層の位置  $i$  から  $m$  へは数密度  $\rho_i$  でシナプス結合があり, 各シナプス結合の荷重は平均すると  $w_i$  であるとする.  $\rho_i$  は固定されていて,  $w_i$  が学習で変化する. L 層の位置  $i$  にあるニューロンの出力を  $x_i$  とすると,  $m$  の出力  $x_m$  はそれらの線形加重和の形になる.

$$x_m = R_a + R_b \sum_i w_i \rho_i x_i \quad (5.1)$$

ここで,  $R_a, R_b$  は定数である.

シナプス結合荷重  $w_i$  は Hebb 則にしたがって学習する.

$$\Delta w_i = k_a + k_b(x_i - x_L)(x_m - x_M) \quad (5.2)$$

$k_a, k_b, x_L, x_M$  は定数である.

### 5.2.2 平均学習方程式

シナプス結合荷重の平均学習方程式を, 式 (5.1) (5.2) から以下のように求めることができる.

$$\dot{w}_i = k_1 + \frac{1}{N_M} \sum_j (Q_{ij} + k_2) \rho_j w_j \quad (5.3)$$

ただし  $N_M = \sum_i \rho_i$  はニューロン  $m$  がうけとる全シナプス数,  $k_1 = (k_a + k_b(\bar{x}_L - x_L)(R_a - x_M))/N_M R_b k_b$ ,  $k_2 = \bar{x}_L(\bar{x}_L - x_L)$  であり,  $\bar{x}_L$  は L 層の各ニューロンの平均出力である.  $Q_{ij}$  は L 層の位置  $i$  のニューロンの出力と位置  $j$  のニューロンの出力との共分散である.

式 (5.3) がシナプス結合荷重  $w_i$  の学習のダイナミクスをあらわしている. このダイナミクスは線形なので一般には  $w_i$  は学習が進むにつれて発散する. そのために, Hebb 則をつかう場合にはふつう,  $w_i$  に対して何らかの制約をつける. Linsker [58-60] では,

ハードリミット形の制限を  $w_i$  につけている。つまり、 $w_i$  が条件  $w_{\min} \leq w_i \leq w_{\max}$  をみたすという制限である。このような制限によって学習の振る舞いは複雑になるが、学習のだいたいの傾向は  $w_i$  が制限の影響をうける以前にほぼ決まってしまう、と考えれば、式 (5.3) を解析することで Linsker のモデルに関するおおよその傾向は知ることができると思われる。

### 5.3 on-中心形 / off-中心形ニューロンの形成

on-中心形 / off-中心形ニューロンの形成に関して、MacKay と Miller [64] が行なった解析について述べる。彼らは Gauss 形の共分散  $Q(r, s) = \exp[-|r - s|^2/2C]$  で、やはり Gauss 形のシナプス結合密度  $\rho(r) = \exp[-|r|^2/2A]$  の場合について解析している。ここで、 $C$ 、 $A$  はモデルのパラメータである。

$k_2 = 0$  の場合をまず考える。便宜上、 $(1/N_M) \sum_j Q_{ij} \rho_j w_j$  を  $w$  に関する線形作用素と考え、 $Q[w]$  とあらわす。モデルのふるまいを解析するためには、 $Q$  の固有値、固有ベクトルを求めればよい。正で最大の固有値に対応する固有ベクトルが実現される傾向にある、ということができるからである。この問題の場合には、共分散の形が原点に関して点対称であるために、群論でよくでてくるように、 $Q$  の固有ベクトルにもそれなりの対称性が要求される。具体的には、固有ベクトルを極座標系であらわしたときに、その偏角成分は  $m$  を整数として  $\cos m\theta$ 、 $\sin m\theta$  といった形をしていなければならない。この条件から  $Q$  の主要な固有ベクトルと、対応する固有値とを求めることができる。これらを表 5.1 に示す。説明の都合上、各固有ベクトルの同心円状の節線、放射状の節線それぞれの数を使って、「(10) モード」などと各固有ベクトルを表現することにする。表 1 に示してあるのは、実際に計算して求めた結果であるが、このほかに  $(n0)$  モードの固有値は  $l^{2n}(1-l)^2$  であることがわかっている。さらに、一般に  $(nm)$  モードの固有値は  $l^{2n+m}(1-l)^2$  であることが推測できる。

この場合には、最大固有値に対応するのが  $(00)$  モードだから、一様な符号をもったシナプス結合が形成される傾向にあるといえる。

つぎに  $k_2 \neq 0$  の場合を考える。 $\rho_i$  の形も空間的には原点に関して点対称だから、 $(01)$  モード、 $(11)$  モードなど定数でない偏角成分をもつ固有ベクトルに関しては、 $k_2$  の値



による影響はない。反対に、(00) モード、(10) モードなどは  $k_2$  によって影響を受ける。各モードの固有値が  $k_2$  の値によってどう変化するか、という問題に関する MacKay と Miller の結果を図 5.2 に示す。

$k_2$  を大きい負の値にとると、(00) モードに代わって (01) モードが最大固有値をもつようになり、形成されるシナプス結合の対称性が破れる。 $k_2$  を 0 に近づけていくと、再び (00) モードが最大固有値をもつようになるが、そこでは  $k_2$  の影響で (00) モードは同心円状の節線をもち、on-中心形 / off-中心形ニューロンができる。 $k_2$  の大きさをもっと小さくしたり、正の値にすると、(00) モードのこの節線が消えてしまい、all-on / all-off 形となる。

$k_1$  の存在は、式 (3) の平衡点を (-0) モードの方向へ移動させることによって、(-0) モードの形成に有利にはたらく。

#### 5.4 方位選択性ニューロンの形成

Linsker [59] は、3 層目に on-中心形 / off-中心形ニューロンができあがると、ニューロン出力の共分散  $Q$  が「メキシカンハット形」でよく近似できるような形になる、としている。そして Linsker の実際のシミュレーションでも、「メキシカンハット形」に理想化した共分散が使われている。それを使って、さらに上位の層で方位選択性ニューロンが形成されているシミュレーション結果が示されている。「メキシカンハット形」の共分散からの方位選択性ニューロンの形成に関して、前節の on-中心形 / off-中心形ニューロンの形成と同様な解析を試み、現在までに得た結果を述べる。

共分散を以下の形にとる。

$$Q(r, s) = (1 - k|r - s|^2)e^{-|r-s|^2/2C} \quad (5.4)$$

パラメータ  $k$  の値が大きいほど、 $Q$  ははっきりしたメキシカンハット形になる (図 5.3)。 $\rho$  は前節と同じものを使う。式 (5.3) での  $k_1$  は 0 とする。式 (5.4) の共分散に対する固有ベクトルは、おそらく前節とことなり解析的に閉じた形で求められない。そこで本論文では、 $k_2$ 、 $k$  のそれぞれに関する一次摂動の範囲で各固有ベクトルに対応する固有

値が、 $k_2$ ,  $k$  によってどのように変化するかを調べて議論する。結果を表 5.2 に示す。表 5.2 での固有関数は、無摂動の形を示してある。

$k_2 = 0$  の場合を考える。 $k$  が小さいとき、つまり共分散  $Q$  の「メキシカンハット度」が小さいときには、 $l$  の値にかかわらず (00) モードが最大固有値をもつ。 $k$  を大きくしていくと、(00) モードの固有値は減少する。他のモードの固有値は、 $l$  の値によって増加したり減少したりする。ここでは、最大固有値をもつのがどのモードであるかに興味があるので、パラメータのそれぞれの値に対して、どのモードが最大固有値をもつかを調べる。結果を  $(Ck)$ - $(C/A)$  平面に図示したものを、図 5.4 に示す。

$Ck < 1/3$  では、 $C/A$  の値によらず (00) モードが最大固有値をもつが、 $Ck$  の値が  $1/3$  をこえると、やがて (01) モードの固有値が (00) モードの固有値より大きくなる。 $Ck$  の値が大きくなると、それにつれてより高次のモードが最大固有値をもつようになる。図に斜線で示した部分は、(00) モードの固有値がもはや正でなくなる領域である。方位選択性ニューロンの性質は、モードでいえば (02) モードに対応していると考えられる。したがってこの結果は、適当なパラメータ値の組み合わせをとれば、Linsker のモデルで方位選択性ニューロンや、さらにはより複雑な刺激選択性をもつニューロンが形成されることをあらわしている。

## 5.5 考察

この章では、初期視覚系での構造の局所的な特徴が自己組織的に形成されることに関するモデルとして Linsker のモデルをとりあげ、とくに Linsker のモデルで方位選択性ニューロンが自己組織的に形成される現象について、解析を試みた。Linsker のモデルは、Hebb 則にしたがう多層フィードバック形の神経場モデルである。Linsker [58-60] は、このモデルの入力層に空間的に無相関な Gauss 形雑音を入力して、Hebb 則にしたがう学習を行なうと、3 層目には on-中心形 / off-中心形ニューロンが、そして早ければ 4 層目には方位選択性ニューロンが自己組織的に形成されることを、計算機シミュレーションの結果として示した。このうち、3 層目での on-中心形 / off-中心形ニューロンの形成に関しては、すでに MacKay と Miller による解析 [64] が行われている。本論

文では、4 層目での方位選択性ニューロンの形成に関して同様の解析を試み、Linsker のモデルで方位選択性ニューロンが自己組織的に形成される可能性を理論的に示した。

この章で得た結果が、実際の生体の神経系での自己組織的形成に関してどのような意味をもつかを考察する。Linsker のモデルによって、on-中心形 / off-中心形ニューロンや方位選択性ニューロンといった、刺激選択性をもつニューロンができあがることが示されたからといって、生体の神経系でも同じ機構によってこれらの刺激選択性ニューロンが形成されているといえるわけではない。いうまでもなく、生体の神経系はここで取り扱ったような Linsker のモデルと比較して桁違いに複雑である。例えば、Linsker のモデルでは、層状フィードフォワード形の構成をもとに議論をしているが、第 2 章で述べたように、生体の神経系にはフィードバックを形成するような逆行性の結合が多数あり、また層内でのニューロンどうしの結合も複雑なものである。また、Linsker のモデルでは興奮性結合と抑制性結合とが同一の枠組みで対等に扱われているが、神経生理学の立場からはこれらの結合様式は互いに異なっているとされており、したがって機能的にも相違があると考えられている。このように考えると、Linsker のモデルが実際の生体の神経系を過度に単純化しているおそれがあることは否定できない。

そのようなことよりもむしろ、Linsker [57] がいうように、Linsker のモデルのような単純な多層神経回路網によって統計的に単純な入力からこのように複雑な性質をもつニューロンが自己組織的に形成されるということ、またこれらのニューロンの性質が初期視覚系でみられる実際の刺激選択性ニューロンの性質と定性的に類似しているということに注目すべきである。さらに、入力がなんら特別な構造をもたなくても、方位選択性ニューロンが自己組織的に形成されて少しもおかしくはない、ということも、本章での結果が示す重要な点である。

実際の初期視覚系で、方位選択性ニューロンにみられる方位に対する選択性が何に由来するか、という問題については、方位選択性ニューロンを発見した当の Hubel と Wiesel が立てた仮説が一般的には受け入れられている。Hubel と Wiesel の仮説とは、一次視覚野の方位選択性ニューロンが入力をうけるニューロンは、LGN のなかでは細長い形をなして分布していて、その方向と方位選択性とが一致している、というものである。このような「選択的に形成される結合」によって方位選択性を説明しようという

立場に対しては、そのような都合のよい結合が選択的に形成されるとは考えにくい、という批判がなされてきた。けれども、この章で得られた結果によれば、なんら複雑な機構を想定しないでも、単純な Hebb 則によるだけで、実際にこのような都合のよい結合が形成されうることになる。また、Stryker ら [9] は、神経生理学的な実験によって、このような「都合のよい結合」が実際に存在していることを示す結果を得ている。

ある霊長類では、生まれたときにすでに、すなわち構造的な視覚経験がなくても、方位選択性ニューロンが形成されることがわかっている。また、LGN や一次視覚野には、左目からあるいは右目からの刺激に対して選択性をもつニューロンが、分離した形で存在している。とくに、左目からあるいは右目からの刺激に対して選択性をもつニューロンは、第 2 章で述べたように LGN では層状構造をつくって分離している。このような構造もやはり、構造的な視覚経験がないうちに形成されることが明らかにされている。またこの構造の形成にあたっては、形成時期に一致して網膜神経節細胞が空間的にランダムな（そしておそらくカオス的な）活動パターンを示すこと、および Hebb 則にしたがうシナプスの存在が関与していると考えられている [79]。この章での結果によれば、一次視覚野での方位選択性ニューロンの形成も、ニューロンの空間的にランダムな活動と Hebb 則とによって理解することができると思われる。

## 第 5 章 図表

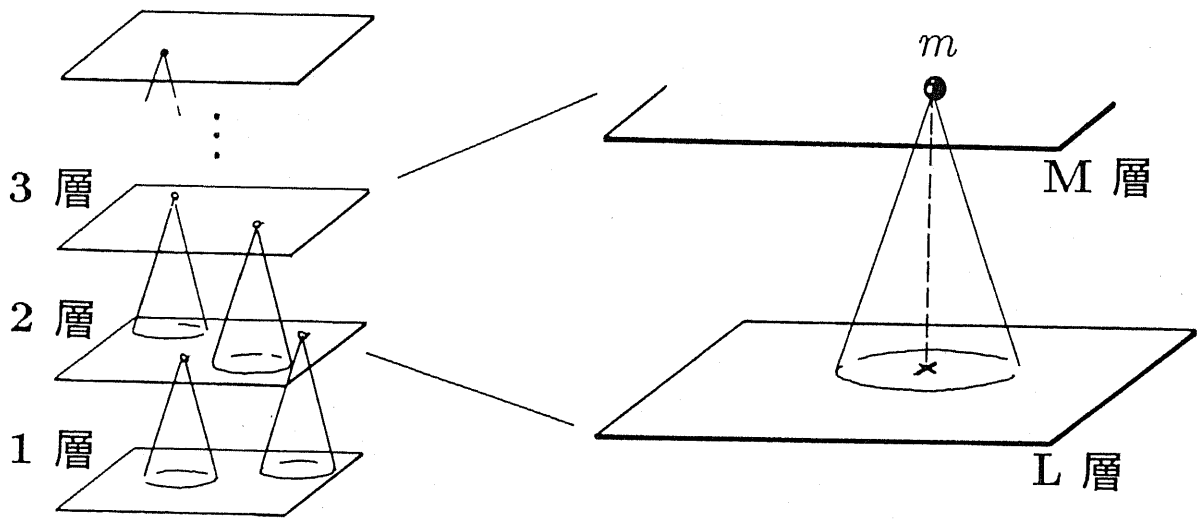


図 5.1 Linsker のモデル

表 5.1  $Q$  の固有ベクトルと固有値

モード	固有関数	固有値
(00)	$e^{-r^2/2R}$	$(1-l)^2$
(01)	$r \cos \theta e^{-r^2/2R}$	$l(1-l)^2$
(10)	$(1 - r^2/r_{(10)}^2) e^{-r^2/2R}$	$l^2(1-l)^2$
(02)	$r^2 \cos 2\theta e^{-r^2/2R}$	$l^2(1-l)^2$
(11)	$r \cos \theta (1 - r^2/r_{(11)}^2) e^{-r^2/2R}$	$l^3(1-l)^2$
(03)	$r^3 \cos 3\theta e^{-r^2/2R}$	$l^3(1-l)^2$
(20)	$(1 - r^2/r_{(20),1}^2 + r^4/r_{(20),2}^4) e^{-r^2/2R}$	$l^4(1-l)^2$
(04)	$r^4 \cos 4\theta e^{-r^2/2R}$	$l^4(1-l)^2$

$$R = \frac{C}{2} \left( 1 + \sqrt{1 + 4A/C} \right)$$

$$l = 1 - \frac{C}{R}$$

$$r_{(10)}^2 = 2A \frac{1+l}{1-l}$$

$$r_{(11)}^2 = 3A \frac{1+l}{1-l}$$

$$r_{(20),1}^2 = \frac{2A(1-l)(2+l^2)}{3(1+l)(1+l^2)}$$

$$r_{(20),2}^4 = \frac{8A^2(1-l)^2(2+l^2)}{(1+l)^2(1+l^2)}$$

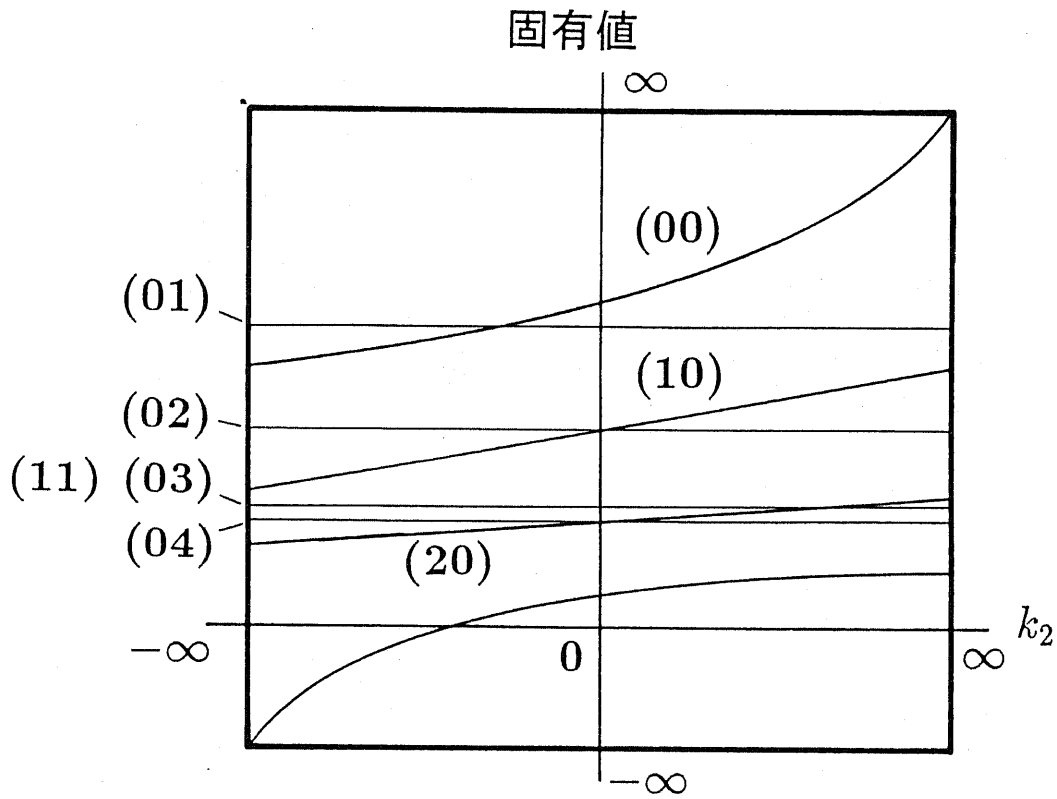
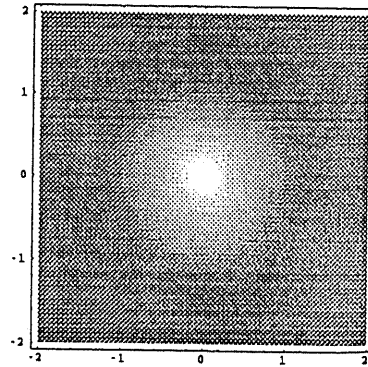
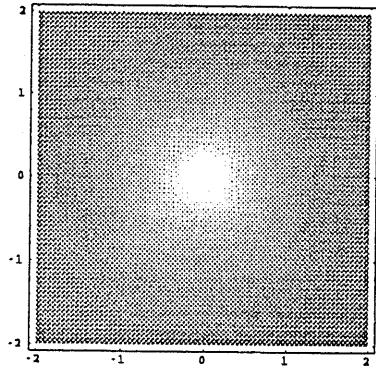
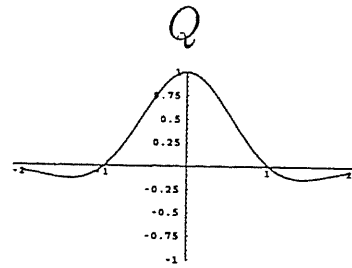
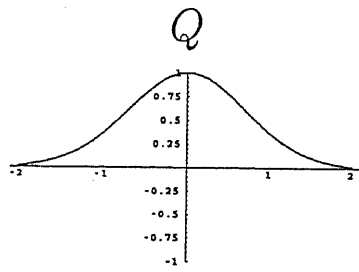


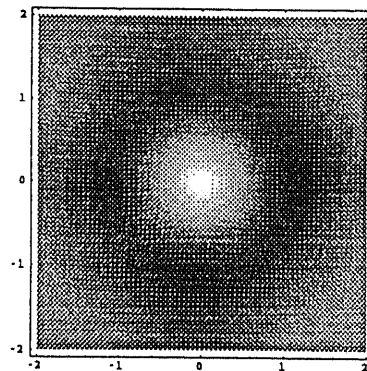
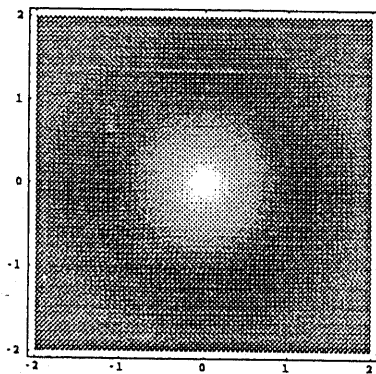
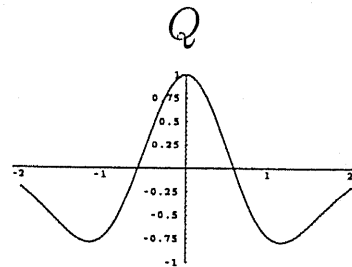
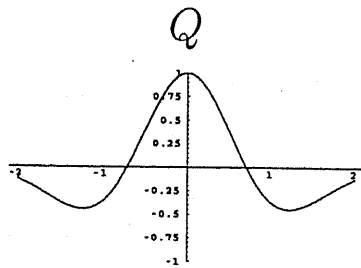
図 5.2  $k_2$ による各固有値の変化 (文献 [64] より)





$Ck = 0$

$Ck = 0.5$



$Ck = 1.0$

$Ck = 1.5$

図 5.3  $Ck$  の値を変化させたときの共分散  $Q$  の形状

表 5.2 メキシカンハット形の共分散に対する固有ベクトルと固有値

モード	固有関数	固有値
(00)	$e^{-r^2/2R}$	$(1-l)^2 \left[ 1 + \frac{1+l}{1-l} k_2 - \frac{4l}{1+l} Ck \right]$
(01)	$r \cos \theta e^{-r^2/2R}$	$l(1-l)^2 \left[ 1 - \frac{2(3l-1)}{1+l} Ck \right]$
(10)	$\left(1 - \frac{1-l^2}{2lC} r^2\right) e^{-r^2/2R}$	$l^2(1-l)^2 \left[ 1 + \frac{1+l}{1-l} k_2 - \frac{4(2l-1)}{1+l} Ck \right]$
(02)	$r^2 \cos 2\theta e^{-r^2/2R}$	$l^2(1-l)^2 \left[ 1 - \frac{4(2l-1)}{1+l} Ck \right]$

$$R = \frac{C}{2} \left( 1 + \sqrt{1 + 4A/C} \right)$$

$$l = 1 - \frac{C}{R}$$

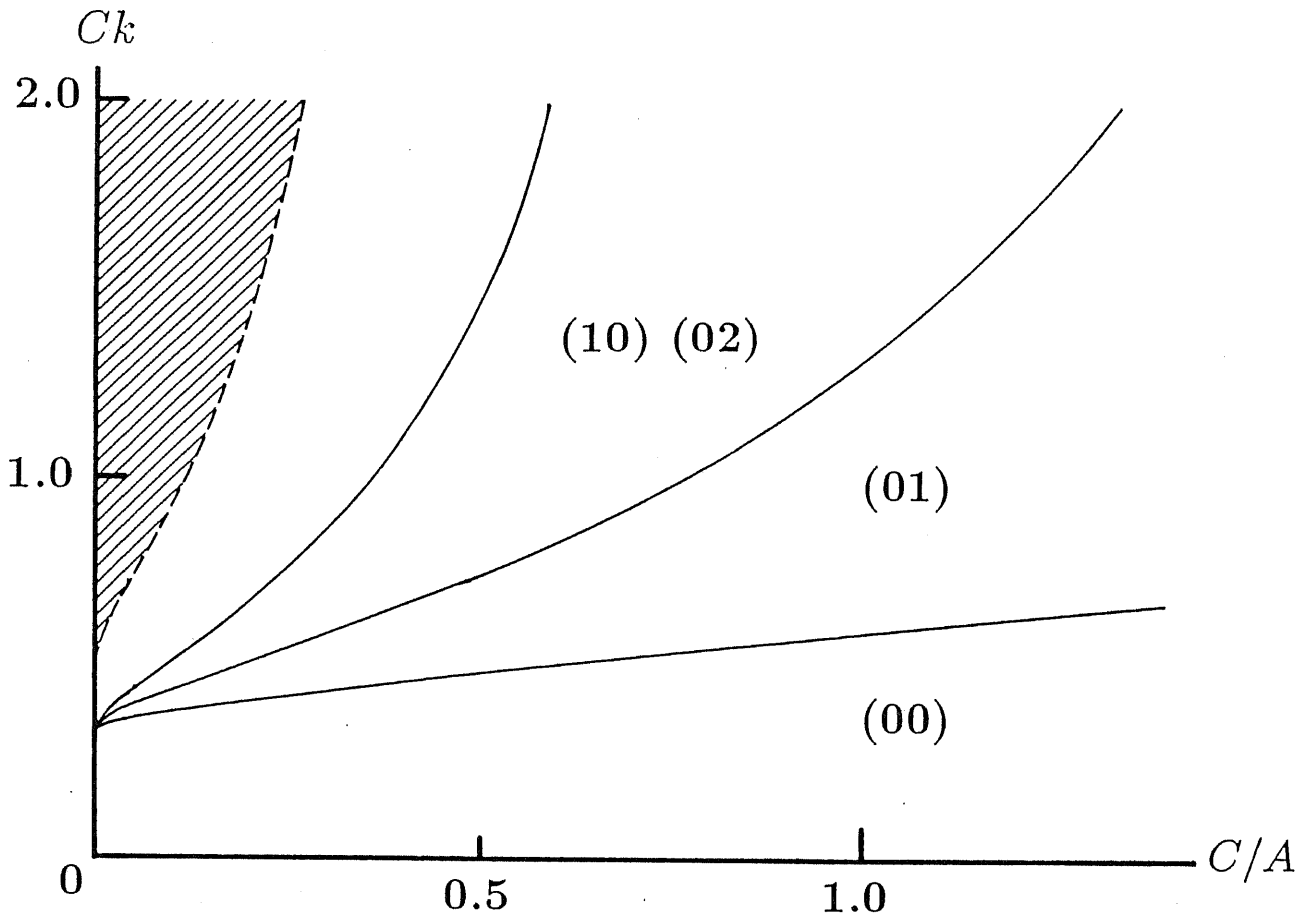


図 5.4 最大固有値をもつモード

## 第 6 章

# 神経系の自己組織的形成に関する考察

猶おもふことひたぶるなるときは言たらず，言  
したらねば思ふ事を末にいひ，仇し語を本に冠  
らせついで，彼五つがひのすがたをたらはせる  
あり ([41])

本論文では、生体の神経系が自己組織的に形成される現象に関して、神経回路網理論の立場から理論的な解析をおこなった。第 4 章では、感覚神経系にみられる位相保存写像のような大域的な構造が、自己組織的に形成されることに関連する、Kohonen のモデルをとりあげて解析した。また、第 5 章では刺激選択性ニューロンのような、感覚神経系にみられる局所的な構造が自己組織的に形成されることに関連して、Linsker のモデルをとりあげて解析した。本論文でとりあげた Kohonen のモデルおよび Linsker のモデルは、どちらも Hebb 則を基本とするモデルである。Hebb 則は、内容が単純で提案された時期こそ古いものの、定性的に Hebb 則にしたがうようなシナプスが生体の神経系に実際に見出されてきたことなどによって、現在でも神経系の学習や自己組織的形成において最も基本的な機構の一つであるとされている。本論文では、Hebb 則にもとづく Kohonen のモデルや Linsker のモデルによって、感覚神経系にみられるような位相保存写像、拡大率に関する比例則や、刺激選択性ニューロンが自己組織的に形成される可能性を示した。また、やはり Hebb 則にもとづくモデルを使い、熱力学におけるスピン系とのアナロジーによって、眼優位性コラムの形成が論じられている研究 [85-87] がある。これらの結果からさらに、神経系のより高次の部分でみられるような構造の自己組織的形成に関しても、Hebb 則にもとづいた、単純な構成規則によって理解することができるのではないか、ということが考えられる。

本論文で議論の対象とした初期視覚系では、線分の方位といった、いわば低い水準の情報が抽出され、処理されている。視覚系は感覚神経系のなかでもとくに高度に発達した情報処理系であり、視覚系のより高次の部分を調べていくと、視覚刺激のさらに複雑な特徴に対して選択性をもつニューロンがある。そのような刺激選択性ニューロンとして、例えば線分の運動方向に選択性をもつニューロンや、パターンやテクスチャの運動方向に選択性をもつもの (サルでは MT 野というところにある [78])、視野全体にわたるほどの大域的な移動、拡大、縮小に選択性をもつもの (MST 野 [78])、色選択性をもつもの (V4 野 [98])、図形特徴に対して選択性をもつもの (TE 野 [84]) などが知られている。顔、または手だけに対して強く反応するようなニューロン (顔ニューロン、手ニューロンなどと呼ばれる [18]) もも報告されている。これらのニューロンでみられ

るような高次の選択性についても、これを実現するための神経系の構造が、なんらかの機構を通して実現されなければならないのはいままでのない。

この章では、MST 野背側部にみられる、視野運動選択性ニューロンを例にあげて考察する。MST 野背側部には、大きな受容野（視角にして平均 40°）をもつ運動方向選択性ニューロンがある。MST 野にみられるニューロンのうち、主なものは以下の 3 種類である。視線に垂直な平面内で並進運動をしたときに方向選択的に反応するニューロン（D ニューロン）、拡大または縮小に選択的に反応するニューロン（E ニューロン、または C ニューロン）、視線に垂直な平面内での、時計回りあるいは反時計回りの回転に選択的に反応するニューロン（Rf ニューロン）。

MST 野背側部の D、E、C、Rf ニューロンは、ほとんどが局所的な運動に対しては感度が低い。反対に、視野の広い範囲にわたる運動に対して最大の反応を示す。それだけでなく、反応の大きさは運動するテクスチャの形にはよらないことが知られている。これらのニューロンが選択的に反応を示すような視覚刺激の特徴は、視覚の対象物が動いたときというよりむしろ、外界に対して目や体が動いたときに生じるものである。いっぽうで、このような大域的な視野運動選択性は、MST 野の前段にあたる MT 野で検出されているような、パターンやテクスチャの局所的な運動情報をうまく統合することで実現されていると考えられている。けれどもそのためには、例えば MST 野で視野全体が上向きに運動したことを検出するニューロンは、MT 野全域に散在する、局所的な上向き運動に選択性をもつニューロンから信号を受けるように、さらに他の方向の運動に選択性をもつような MT 野の大多数のニューロンからの信号はなるべく受けないように、結合を調整しなければならない。このような都合のよい結合様式は、どのような機構によって形成されるのだろうか。

この問題に対して、本論文の立場から提案される仮説は、「そのような都合のよい結合を形成するには、Hebb 則のような原理が有効にはたらいっている」というものである。いまかりに、MT 野で遠く離れている 2 つのニューロンが、どちらも局所的な上向き運動に対する選択性をもっていて、MST 野のあるニューロンがごくわずかに大域的な上向き運動に対する選択性をもったとしよう。すると、視野全体にわたる上向き運動が生じるたびに、MT 野の 2 つのニューロンから MST 野のこのニューロンへの結合

は、それらがいかに空間的に離れていようとも、Hebb 則によって常に同時に強化される。生後の視覚環境を考えれば、視野運動はきわめて高い頻度でおこる入力刺激だから、MT 野に形成されるような局所的な運動方向選択性をもつニューロンを前提にすれば、視野運動選択性をもつニューロンは、このような機構によって容易に形成されると考えられる。

もしこの仮説が妥当なものであるなら、MST 野で視野運動選択性ニューロンが形成されるためには、Hebb 則を実現するような機構に加えて、局所的な運動方向選択性ニューロンが形成される脳の領野の全域から入力を受けるように、MST 野のニューロンが遺伝的あるいは形態学的に条件づけられていけばよい、ということになる。大脳皮質における機能地図が個体によって大きく異なったりしない、という事実と併せて考えると、大脳皮質の機能分化は形態学的な要因によって強く条件づけられている、ということはこの仮説は意味しているように思われる。このことを確認する、あるいは否定するためには、大脳生理学、神経科学などからの実験的、理論的な知見の集積が必要であろうが、将来的には検討されるべき課題であると考えている。

## 第 7 章

### 結論

まことに、われもひとも、そらごとをのみまう  
しあひさふらふなかに、ひとついたましきこと  
のさふらふなり。 ([100])



本論文では、生体の神経系の自己組織的形成について、神経回路網理論の立場から提案されているモデルを使って理論的な解析をおこなった。おもに初期視覚系の自己組織的形成を対象とし、神経回路網理論の立場からのモデルとして、生理学的に妥当性があり単純な Hebb 則にもとづく、Kohonen のモデル、および Linsker のモデルをとりあげた。

Kohonen のモデルは、初期視覚系でみられる構造の大域的な特徴の自己組織的形成に関するモデルとしてとりあげた。Kohonen のモデルに関しては、モデルによって形成される写像の拡大率と入力信号の頻度との関係を定量的に議論した。粗視的にみれば、写像の拡大率が入力信号の頻度のべきに比例する形へ、安定に収束することを理論的に示した。この結果から、生体の感覚神経系でいわれている、写像の拡大率に関する比例則が、Kohonen のモデルでも定性的に再現されることを示した。

Kohonen のモデルに関する以上の理論的解析を、モデルの次元を限定しないで行ったので、生体の神経系に関連する 1, 2 次元の場合だけでなく、適応的信号処理システムとして Kohonen のモデルを応用する場合にも適用できる一般的な形で結果を得た。

また、Kohonen のモデルによる位相保存写像の形成に関して議論し、モデルが 1 次元の場合には、十分長い時間をかければ位相保存写像が確率 1 で形成される、という、一般的な結果を証明した。

Linsker のモデルは、初期視覚系でみられる構造の局所的な特徴、具体的には刺激選択性ニューロンの、自己組織的形成に関するモデルとしてとりあげた。Linsker のモデルに関しては、早ければモデルの第 4 層で方位選択性ニューロンが形成されるという、計算機シミュレーションによる実験的な観察に対し、摂動計算による理論的な解析をおこなった。第 4 層への入力の間接性がメキシカンハット形となることにより、モデルの回転対称性が自発的に破れ、方位選択性ニューロンが形成される可能性を、理論的に提示した。

さらに、視覚系でより高次の刺激選択性を示すニューロンを例にとりあげ、その形成も Hebb 則で理解できることを示した。

以上の結果から、生体の神経系の自己組織的形成において Hebb 則が基本的な重要であること、またその解析に神経回路網理論の立場からのモデルとその数理的な解析との有効性を示した。以上が本論文の結論である。

## 謝辞

本研究は、東京大学医学部医用電子研究施設 斎藤正男教授 (東京大学大学院工学系研究科電子工学専攻教授を併任) の指導のもとで行なわれたものである。浅学非才の筆者がこのような身分不相応な研究を遂行できたのも、同教授の懐の深い指導のおかげであると思っている。

本論文をまとめるにあたって、新技術事業団さきがけ研究 21 および理化学研究所国際フロンティア研究システム思考電流チーム 藤田一郎氏には、神経生理学の分野の文献を多数紹介していただいた。また同氏が 1992 年に東京大学工学部計数工学科でおこなった講義は、筆者にとって非常に有益なものであった。

文献および基礎的な資料の調査にあたり、研究室の大学院生諸氏ならびに田中義浩氏、山家菊世氏に御協力いただいた。

本研究の一部は、文部省科学研究費補助金 (一般研究 B, 課題番号 03452183) の援助をうけておこなわれた。

本論文は、文書組版システム  $\text{\LaTeX}$  によって構成されており、 $\text{\TeX}$ ,  $\text{\LaTeX}$  およびそれに関連する多くの無償のソフトウェアを利用した。 $\text{\TeX}$  およびそれに関連するソフトウェアに関して、倉島顕尚氏に有益な助言をいただいた。本研究を進めるに際して、東京大学大型計算機センターを利用した。

## 参考文献

- [1] S. C. Ahalt, A. K. Krishnamurthy, P. Chen and D. E. Melton, "Competitive learning algorithms for vector quantization," *Neural Networks*, vol. 3, no. 3, pp. 277-290, 1990.
- [2] 合原一幸, ニューラルコンピュータ: 脳と神経に学ぶ, 東京電機大学出版局, 1988.
- [3] Shun-ichi Amari, "Topographic organization of nerve fields," *Bulletin of Mathematical Biology*, vol. 42, pp. 339-364, 1980.
- [4] M. A. Arbib, *The Metaphorical Brain 2: Neural Networks and Beyond*, Wiley-Interscience, New York, 1989, (金子隆芳 訳, ニューラルネットと脳理論—アービブ『脳』第2版—, サイエンス社, 1992).
- [5] Robert Azencott, editor, *Simulated Annealing: Parallelization Techniques*, A Wiley-Interscience Publication, John Wiley & Sons, New York, Chichester, Brisbane, Toronto, Singapore, 1992.
- [6] Hans-Ulrich Bauer and Klaus R. Pawelzik, "Quantifying the neighborhood preservation of self-organizing feature maps," *IEEE Trans. Neural Networks*, vol. 3, no. 4, pp. 570-579, July 1992.
- [7] Edward K. Blum and Leong Kwan Li, "Approximation theory and feedforward networks," *Neural Networks*, vol. 4, no. 4, pp. 511-515, 1991.
- [8] Pierre Cardaliaguet and Guillaume Euvrard, "Approximation of a function and its derivative with a neural network," *Neural Networks*, vol. 5, no. 2, pp. 207-220, 1992.

- [9] Barbara Chapman, Kathleen R. Zahs and Michael P. Stryker, "Relation of cortical cell orientation selectivity to alignment of receptive fields of the geniculocortical afferents that arborize within a single orientation column in ferret visual cortex," *The Journal of Neuroscience*, vol. 11, no. 5, pp. 1347-1358, May 1991.
- [10] Patricia S. Churchland and Terrence J. Sejnowski, *The Computational Brain*, The MIT Press, Cambridge, MA, 1992.
- [11] Marie Cottrell et Jean-Claude Fort, "Étude d'un processus d'auto-organisation," *Annales de l'Institut Henri Poincaré - Probabilités et Statistiques*, vol. 23, n<sup>o</sup>. 1, pp. 1-20, 1987.
- [12] P. M. Daniel and D. Whitteridge, "The presentation of the visual field on the cerebral cortex in monkeys," *Journal of Physiology (London)*, vol. 159, no. 2, pp. 203-221, December 1961.
- [13] D. DeSieno, "Adding a conscience to competitive learning," *IEEE Int. Conf. Neural Networks*, vol. 1, pp. I-117-124, 1988.
- [14] Rodney J. Douglas, Kevan A. C. Martin and David Whitteridge, "A canonical microcircuit for neocortex," *Neural Computation*, vol. 1, no. 4, pp. 480-488, 1989.
- [15] Richard O. Duda and Peter E. Hart, *Pattern classification and scene analysis*, chapter 6, A Wiley-Interscience Publication, John Wiley & Sons, New York, London, Sydney, Toronto, 1973.
- [16] E. Erwin, K. Obermayer and K. Schulten, "Self-organizing maps: stationary states, metastability and convergence rate," *Biological Cybernetics*, vol. 67, no. 1, pp. 35-45, May 1992.
- [17] E. Erwin, K. Obermayer and K. Schulten, "Self-organizing maps: ordering, convergence properties and energy functions," *Biological Cybernetics*, vol. 67, no. 1, pp. 47-55, May 1992.

- [18] 藤田一郎, 下側頭回の形態識別機能, 神経進歩, vol. 35, no. 3, pp. 414-422, June 1991.
- [19] 藤田一郎, 耳のずれたフクロウ: 聴覚空間認識の脳内機構, 生物物理, vol. 32, no. 3, pp. 147-153, March 1992.
- [20] 船橋賢一, ニューラル・ネットワークによる連続写像の近似的実現について, 信学技報, MBE 88-9, 1988.
- [21] 船橋賢一, ニューラル・ネットワークの capability について, 信学技報, MBE 88-52, 1988.
- [22] Ken-ichi Funahashi, "On the approximate realization of continuous mappings by neural networks," *Neural Networks*, vol. 2, pp. 183-192, 1989.
- [23] 船橋賢一, 3層ニューラルネットワークによる恒等写像の近似的実現についての理論的考察, 電子情報通信学会論文誌 A, vol. J73-A, no. 1, pp. 139-145, January 1990.
- [24] Stuart Geman, "Some averaging and stability results for random differential equations," *SIAM J. Appl. Math.*, vol. 36, no. 1, pp. 86-105, February 1979.
- [25] Stuart Geman and Donald Geman, "Stochastic relaxation, Gibbs distribution, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-6, no. 6, pp. 721-741, November 1984, (柳研二郎 訳, 統計的緩和法, Gibbs 分布と画像のベイズ修復, 甘利 俊一, 後藤英一 編, 人工ニューラルシステム, bit 臨時増刊, vol. 21, no. 11, pp. 1442-1466, 1989年9月).
- [26] Federico Girosi and Tomaso Poggio, "Representation properties of networks: Kolmogorov's theorem is irrelevant," *Neural Computation*, vol. 1, no. 4, pp. 465-469, 1989.

- [27] Hermann Haken, *Advanced Synergetics: Instability Hierarchies of Self-Organizing Systems and Devices*, Springer-Verlag, 1983, (斎藤 信彦, 小森尚志, 長島 知正 訳, シナジェティクスの基礎: 不安定性の階層=システムとデバイスの自己組織化, 東海大学出版会, 1986).
- [28] Donald O. Hebb, *The Organization of Behavior*, chapter 4, p. 62, Wiley, New York, 1949, reprinted in *Neurocomputing: Foundations of Research*, James A. Anderson and Edward Rosenfeld, eds., The MIT Press, Cambridge, MA, Chapter 4, pp. 45-58, 1988.
- [29] R. Hecht-Nielsen, "Theory of the backpropagation neural network," *Proc. Int. Joint Conf. Neural Networks*, I, pp. 593-611, June 1989.
- [30] Robert Hecht-Nielsen, "Kolmogorov's mapping neural network existence theorem," *Proc. IEEE 1st Int. Conf. Neural Networks*, vol. III, pp. 11-14, 1987.
- [31] Robert Hecht-Nielsen, *Neurocomputing*, Addison-Wesley Publishing Company, 1990, (袋谷賢吉 訳, ニューロコンピューティング, アジソン ウェスレイ・トッパン 情報科学シリーズ, 24, トッパン, 1992).
- [32] Arun V. Holden, *Models of the stochastic activity of neurones*, vol. 12 of *Lecture Notes in Biomathematics*, Springer-Verlag, Berlin, Heidelberg, New York, 1976.
- [33] John J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," *Proceedings of the National Academy of Sciences USA*, vol. 79, pp. 2554-2558, 1982.
- [34] John J. Hopfield, "Neurons with graded response have collective computational properties like those of two-state neurons," *Proceedings of the National Academy of Sciences USA*, vol. 81, pp. 3088-3092, 1984.
- [35] Kurt Hornik, "Approximation capabilities of multilayer feedforward networks," *Neural Networks*, vol. 4, no. 2, pp. 251-257, 1991.

- [36] Kurt Hornik, Maxwell Stinchcombe and Halbert White, "Multilayer feedforward networks are universal approximators," *Neural Networks*, vol. 2, pp. 359-366, 1989.
- [37] D. H. Hubel and T. N. Wiesel, "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex," *Journal of Physiology (London)*, vol. 160, pp. 106-154, 1962.
- [38] M. Iri, K. Murota and T. Ohya, "A fast Voronoi-diagram algorithm with applications to geographical optimization problems," *Proc. 11th IFIP Conf. System Modelling and Optimization 1983, Copenhagen*, Lecture Notes in Control and Information Science, 59, System Modelling and Optimization (P. Thoft-Christensen, ed.), pp. 273-288, Springer-Verlag, 1984.
- [39] Bunpei Irie and Sei Miyake, "Capabilities of three-layered perceptrons," *Proc. IEEE Int. Conf. Neural Networks*, vol. I, pp. 641-648, 1988.
- [40] Yoshifusa Ito, "Representation of functions by superpositions of a step or sigmoid function and their applications to neural network theory," *Neural Networks*, vol. 4, no. 3, pp. 385-394, 1991.
- [41] 賀茂真淵, 冠辭考 序文.
- [42] Eric R. Kandel and Robert D Hawkins, "The biological basis of learning and individuality," *Scientific American*, vol. 267, no. 3, pp. 53-60, September 1992, (津本 忠治, 田村 弘 訳, ニューロンレベルでみた学習, 日経サイエンス, vol. 22, no. 11, pp. 50-61, 1992年11月).
- [43] 柄谷行人, 心理を超えたものの影, 畏怖する人間, 冬樹社, 1972.
- [44] 柄谷行人, 言語・数・貨幣, 内省と廻行, 講談社, 1985.
- [45] 川人光男, 乾敏郎, 視覚大脳皮質の計算理論, 電子情報通信学会論文誌, D-II, vol. J73-D-II, no. 8, pp. 1111-1121, August 1990.

- [46] 川人光男, 乾敏郎, 視覚皮質の計算理論: 再構成からパターン認識, 記憶まで, 科学, vol. 61, no. 4, pp. 214-222, April 1991.
- [47] S. Kirkpatrick, Jr. C. D. Gelatt and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, no. 4598, pp. 671-681, 13 May 1983.
- [48] Christof Koch and Heinz Shuster, "A simple network showing burst synchronization without frequency locking," *Neural Computation*, vol. 4, no. 2, pp. 211-223, March 1992.
- [49] Teuvo Kohonen, *Self-Organization and Associative Memory*, third edition, Springer Series in Information Sciences, 8, Springer-Verlag, Berlin, Heidelberg, New York, London, Paris, Tokyo, Hong Kong, 1989.
- [50] 倉田耕治, 競合的な隠れユニットを持つ神経回路網の学習について, 数理解析研究所講究録, 第 678 巻, pp. 52-67, 京都大学数理解析研究所, 1989.
- [51] 倉田耕治, トポグラフィックマッピング自己組織モデルにおけるコラムおよびハイパーコラムの形成, 信学技報, NC89-72, March 1990.
- [52] 倉田耕治, 山田徹, 発火強度の変化するトポグラフィック・マッピング形成モデルにおける離散微細構造の形成, 信学技報, NC90-108, March 1991.
- [53] Věra Kůrková, "Kolmogorov's theorem is relevant," *Neural Computation*, vol. 3, pp. 617-622, 1991.
- [54] Věra Kůrková, "Kolmogorov's theorem and multilayer neural networks," *Neural Networks*, vol. 5, no. 3, pp. 501-506, 1992.
- [55] Pierre Landau and Eric Schwartz, "Computer simulation of cortical polymaps: A proto-column algorithm," *Neural Networks*, vol. 5, no. 2, pp. 187-206, 1992.
- [56] Yoseph Linde, Andrés Buzo and Robert M. Gray, "An algorithm for vector quantizer design," *IEEE Trans. Commun.*, vol. COM-28, no. 1, pp. 84-95, January 1980.



- [57] R. Linsker, "Self-organization in a perceptual network," *Computer*, vol. 21, no. 3, pp. 105-111, March 1988.
- [58] Ralph Linsker, "From basic network principles to neural architecture: Emergence of spatial-opponent cells," *Proceedings of the National Academy of Sciences USA*, vol. 83, no. 19, pp. 7508-7512, October 1986.
- [59] Ralph Linsker, "From basic network principles to neural architecture: Emergence of orientation-selective cells," *Proceedings of the National Academy of Sciences USA*, vol. 83, no. 21, pp. 8390-8394, November 1986.
- [60] Ralph Linsker, "From basic network principles to neural architecture: Emergence of orientation columns," *Proceedings of the National Academy of Sciences USA*, vol. 83, no. 22, pp. 8779-8783, November 1986.
- [61] Ralph Linsker, "How to generate ordered maps by maximizing the mutual information between input and output," *Neural Computation*, vol. 1, no. 3, pp. 402-411, 1989.
- [62] Z.-P. Lo and B. Bavarian, "On the rate of convergence in topology preserving neural networks," *Biological Cybernetics*, vol. 65, no. 1, pp. 55-63, May 1991.
- [63] Stephen P. Luttrell, "Code vector density in topographic mappings: Scalar case," *IEEE Trans. Neural Networks*, vol. 2, no. 4, pp. 427-436, July 1991.
- [64] D. J. C. MacKay and K. D. Miller, "Analysis of Linsker's simulations of Hebbian rules," *Neural Computation*, vol. 2, no. 2, pp. 173-187, 1990.
- [65] David Marr, *Vision*, Freeman, New York, 1981, (乾 敏郎, 安藤広志 訳, ビジョン: 視覚の計算理論と脳内表現, 産業図書, 1987).
- [66] Karl Marx, 経済学批判序説, 経済学批判, 岩波文庫, 1956.
- [67] Karl Marx, "Das Kapital, Kritik der politischen ökonomie," MEGA, Bd. II/6, Dietz Verlag, 1987.

- [68] Warren S. McCulloch and Walter H. Pitts, "A logical calculus of ideas immanent in nervous activity," *Bulletin of Mathematical Biophysics*, vol. 5, pp. 115-133, 1943, reprinted in Warren S. McCulloch, *Embodiments of Mind*, The MIT Press, Cambridge, MA, pp. 19-39, 1965; and also in *Neurocomputing: Foundations of Research*, James A. Anderson and Edward Rosenfeld, eds., The MIT Press, Cambridge, MA, Chapter 2, pp. 18-27, 1988.
- [69] Marvin Minsky and Seymour Papert, *Perceptrons*, The MIT Press, 1969, (斎藤正男 訳, パーセプトロン, 東京大学出版会, 1971).
- [70] D. Mumford, "On the computational architecture of the neocortex: I. the role of the thalamo-cortical loop," *Biological Cybernetics*, vol. 65, no. 2, pp. 135-145, June 1991.
- [71] 紫式部, 源氏物語.
- [72] James A. Reggia, C. Lynne D'Autrechy, Granger G. Sutton III and Michael Weinrich, "A competitive distribution theory of neocortical dynamics," *Neural Computation*, vol. 4, no. 3, pp. 287-317, May 1992.
- [73] H. Ritter and K. Schulten, "On the stationary state of Kohonen's self-organizing sensory mapping," *Biological Cybernetics*, vol. 54, pp. 99-106, 1986.
- [74] H. Ritter and K. Schulten, "Convergence properties of Kohonen's topology conserving maps: Fluctuations, stability, and dimension selection," *Biological Cybernetics*, vol. 60, pp. 59-71, 1988.
- [75] Helge Ritter, "Asymptotic level density for a class of vector quantization processes," *IEEE Trans. Neural Networks*, vol. 2, no. 1, pp. 173-175, January 1991.
- [76] Anna W. Roe, Sarah L. Pallas, Jong-On Hahm and Mriganka Sur, "A map of visual space induced in primary auditory cortex," *Science*, vol. 250, no. 4982, pp. 818-820, 9 November 1990.

- [77] D. E. Rumelhart and J. L. McClelland, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, I & II*, The MIT Press, Cambridge MA, 1986, (甘利俊一 監訳, PDP モデル — 認知科学とニューロン回路網の探索一, 産業図書, 1989).
- [78] 斎藤秀昭, 脳における視覚的運動の階層処理, 医学のあゆみ, vol. 161, no. 11, pp. 847-850, June 1992.
- [79] Carla J. Shatz, "The developing brain," *Scientific American*, vol. 267, no. 3, pp. 34-41, September 1992, (武田 泰生, 植村 慶一 訳, 脳の発生過程, 日経サイエンス, vol. 22, no. 11, pp. 26-36, 1992 年 11 月).
- [80] Alcino J. Silva, Richard Paylor, Jeanne M. Wehner and Susumu Tonegawa, "Impaired spatial learning in  $\alpha$ -calcium-calmodulin kinase II mutant mice," *Science*, vol. 257, no. 5067, pp. 206-211, 10 July 1992.
- [81] Alcino J. Silva, Charles F. Stevens, Susumu Tonegawa and Yanyan Wang, "Deficient hippocampal long-term potentiation in  $\alpha$ -calcium-calmodulin kinase II mutant mice," *Science*, vol. 257, no. 5067, pp. 201-206, 10 July 1992.
- [82] William R. Softky and Christof Koch, "Cortical cells should fire regularly, but do not," *Neural Computation*, vol. 4, no. 5, pp. 643-646, September 1992.
- [83] Mriganka Sur, Preston E. Garraghty and Anna W. Roe, "Experimentally induced visual projections into auditory thalamus and cortex," *Science*, vol. 242, no. 4884, pp. 1437-1441, 9 December 1988.
- [84] 田中啓治, 物体像の脳内表現, 科学, vol. 61, no. 4, pp. 231-238, April 1991.
- [85] Shigeru Tanaka, "Theory of self-organization of cortical maps: Mathematical framework," *Neural Networks*, vol. 3, pp. 625-640, 1990.

- [86] Shigeru Tanaka, "Theory of ocular dominance column formation: Mathematical basis and computer simulation," *Biological Cybernetics*, vol. 64, no. 4, pp. 263-272, February 1991.
- [87] Shigeru Tanaka, "Phase transition theory for abnormal ocular dominance column formation," *Biological Cybernetics*, vol. 65, no. 2, pp. 91-98, June 1991.
- [88] 田中利幸, 神経素子の応答の速さについて, 信学技報, vol. MBE90-6, , April 1990.
- [89] 田中利幸, 斎藤正男, 閾値制御による神経素子の入出力静特性について, 信学技報, vol. MBE89-90, , December 1989.
- [90] 田中利幸, 斎藤正男, Kohonen のモデルが実現する写像について, 信学技報, NC90-107, March 1991.
- [91] 田中利幸, 斎藤正男, Kohonen のモデルにおける学習の収束性について, 信学技報, NC91-154, March 1992.
- [92] 田中利幸, 斎藤正男, 適応的ベクトル量子化法としての Kohonen のモデルの定量的性質, 電子情報通信学会論文誌 D-II 分冊, vol. J75-D-II, no. 6, pp. 1085-1092, June 1992.
- [93] V. V. Tolat, "An analysis of Kohonen's self-organizing maps using a system of energy functions," *Biological Cybernetics*, vol. 64, no. 2, pp. 155-164, December 1990.
- [94] Chr. von der Malsburg, "Self-organization of orientation sensitive cells in the striate cortex," *Kybernetik*, bd. 14, heft 2, pp. 85-100, 1973.
- [95] Norbert Wiener, *Cybernetics*, The MIT Press, Cambridge, MA, 1947, (池原止戈夫, 彌永昌吉, 室賀三郎, 戸田巖 共訳, サイバネティックス (第2版), 岩波書店, 1962).
- [96] 山田徹, 倉田耕治, コホネンのトポグラフィック・マッピング自己組織モデルにおける離散微細構造の形成, 信学技報, NC90-109, March 1991.

- [97] Paul L. Zador, "Asymptotic quantization error of continuous signals and the quantization dimension," *IEEE Trans. Inform. Theory*, vol. IT-28, no. 2, pp. 139-149, March 1982, (revised and expanded version of the Bell Laboratories unpublished memorandum "Topics in the asymptotic quantization of continuous random variables," Bell Laboratories, Murray Hill, NJ, February 1966).
- [98] Semir Zeki, "The visual image in mind and brain," *Scientific American*, vol. 267, no. 3, pp. 43-50, September 1992, (赤瀬 英介, 外山敬介 訳, 脳と視覚, 日経サイエンス, vol. 22, no. 11, pp. 38-49, 1992年11月).
- [99] Jun Zhang, "Dynamics and formation of self-organizing maps," *Neural Computation*, vol. 3, no. 1, pp. 54-66, 1991.
- [100] 歎異抄 十九.

## 付録 A

## 式 (4.28) の導出

本文での、式 (4.27) からの式 (4.28) の導出を示す。

式 (4.27) の左辺を

$$I = \int_{C'} |f(z)|^{r-2} f(z) p(\mathbf{w}_i + f(z)) J[f] dz \quad (\text{A.1})$$

とおく。  $|z|$  が十分小さいと考えて、被積分関数のテイラー展開の高次項を無視する近似を行う。まず、

$$f(z) = (1 - ce_g \cdot z)z \quad (\text{A.2})$$

だから、

$$|f(z)|^{r-2} f(z) = (1 - (r-1)ce_g \cdot z) |z|^{r-2} z \quad (\text{A.3})$$

$$J[f] = 1 - (n+1)ce_g \cdot z \quad (\text{A.4})$$

である。また、  $p_0 = p(\mathbf{w}_i)$ ,  $\nabla p_0 = \nabla_{\mathbf{x}} p(\mathbf{x})|_{\mathbf{x}=\mathbf{w}_i}$  とおくと、

$$p(f(z)) = p_0 + (1 - ce_g \cdot z) \nabla p_0 \cdot z \quad (\text{A.5})$$

である。したがって、

$$I = \int_{C'} |z|^{r-2} z [p_0 + \nabla p_0 \cdot z - c(e_g \cdot z) \{(n+r)p_0 + (n+r+1)(\nabla p_0 \cdot z)\}] dz \quad (\text{A.6})$$

となる。

積分領域  $C'$  は  $z=0$  を中心とする超立方体だから、対称性に注意すると、 $a, b$  を定ベクトルとして、以下の関係式を得る。

$$\begin{aligned} \int_{C'} |z|^{r-2} z dz &= 0 \\ \int_{C'} |z|^{r-2} (a \cdot z) z dz &= K a \\ \int_{C'} |z|^{r-2} (a \cdot z) (b \cdot z) z dz &= 0 \end{aligned} \quad (\text{A.7})$$

ここで、 $K$  はスカラーで、

$$K = \int_{C'} |z|^{r-2} z_1^2 dz \quad (\text{A.8})$$

である。

これらを使って式 (A.6) を計算することによって、

$$I = K[\nabla p_0 - c(n+r)p_0 e_g] \quad (\text{A.9})$$

を得る。式 (4.27) は  $I=0$  であったから、式 (A.9) から

$$K[\nabla p_0 - c(n+r)p_0 e_g] = 0 \quad (\text{A.10})$$

が得られる。表記をもとに戻すと、式 (4.28) すなわち

$$K[\nabla p(\mathbf{w}_i) - c(n+r)p(\mathbf{w}_i)e_g] = 0 \quad (\text{A.11})$$

が導かれる.

$f(z)$  は参照ベクトルの配置の一様な配置からのずれを表しているが, ここでは  $f(z)$  は式 (A.2) のような形をしていることを仮定した. 一般には  $f(z)$  は  $z$  のより高次の項を含んでいる. しかしここで行ったような高次項を無視する近似が成立するならば, ここでの結果が  $f(z)$  の高次項によって影響をうけることはない. また, 積分領域  $C'$  はここでは  $z = 0$  を中心とする超立方体を考えているが, 式 (A.7) は  $C'$  の対称性だけからえられるものだから, 例えば  $C'$  が他の超多面体や超球といった領域であったとしても, ここでの結果は全く同様に成立する.



## 付録 B

## 式 (4.38) の導出

本文での式 (4.38) の導出を示す.

まず  $X = q/(cp^s)$  であることより,

$$\begin{aligned} \frac{\nabla X}{X} &= \nabla(\log X) \\ &= \nabla(\log q) - \nabla(\log(cp^s)) \\ &= \frac{\nabla q}{q} - s \frac{\nabla p}{p} \end{aligned} \tag{B.1}$$

である. このことから,

$$\dot{x} = -\frac{K}{s} p \frac{\nabla X}{X} \tag{B.2}$$

とあらわされる.

ここで, 以下の補題を用意しておく.

**補題**  $S$  を  $R^n$  の有界領域,  $\varphi: S \mapsto R$ ,  $A: S \mapsto R^n$  とし,  $S$  の境界  $\partial S$  上で  $\varphi A \equiv 0$  であるとする. このとき,

$$\int_S \varphi \nabla \cdot A \, dx = - \int_S (\nabla \varphi) \cdot A \, dx \tag{B.3}$$

がなりたつ.

**証明** 公式

$$\nabla \cdot (\varphi A) = \varphi \nabla \cdot A + (\nabla \varphi) \cdot A \tag{B.4}$$

の両辺を  $S$  で積分する.

$$\int_S \nabla \cdot (\varphi A) dx = \int_S \varphi \nabla \cdot A dx + \int_S (\nabla \varphi) \cdot A dx \quad (\text{B.5})$$

ここからさらに, Gauss の公式を使って式 (B.5) の左辺を  $\partial S$  上の面積分でかきなおすことができる.

$$\int_S \nabla \cdot (\varphi A) dx = \int_{\partial S} (\varphi A) \cdot n ds \quad (\text{B.6})$$

ここで,  $ds$  は  $\partial S$  上の面積素片,  $n$  は  $ds$  上の外向き単位法線ベクトルである.

仮定から,  $S$  の境界  $\partial S$  上で  $\varphi A \equiv 0$  だから, 式 (B.6) の値は 0 に等しい. したがって式 (B.5) の左辺も 0 に等しいので, このことから証明すべき式 (B.3) を得る. (証明おわり)

$j$  の計算を続ける.

$$\begin{aligned} j &= \int_S \delta J \cdot \dot{q} dx \\ &= - \int_S \frac{1}{q} X f'(X) q \nabla \cdot \dot{x} dx \\ &= - \int_S X f'(X) \nabla \cdot \dot{x} dx \end{aligned} \quad (\text{B.7})$$

$S$  の境界  $\partial S$  上では  $X = 0$  であると考えられるから, ここで付録のはじめに示した補題を使うことができる. その結果として

$$j = \int_S \nabla [X f'(X)] \cdot \dot{x} dx \quad (\text{B.8})$$

を得る.

式 (B.2) を代入して計算を進めることにより, 式 (4.38) に対応する以下の結果を得る.

$$\begin{aligned} j &= -\frac{K}{s} \int_S [X f''(X) + f'(X)] \nabla X \cdot \left( p \frac{\nabla X}{X} \right) dx \\ &= -\frac{K}{s} \int_S \left[ f''(X) + \frac{f'(X)}{X} \right] p \|\nabla X\|^2 dx \end{aligned} \quad (\text{B.9})$$

## 付録 C

### 定理の証明

**定理** 1次元の重みつき学習において、入力信号が有限区間  $[a, b]$  上に発生し、入力信号の確率密度分布がこの区間で 0 になることはないとする。また、参照ベクトル数を  $N$  とし、近傍関数  $h(m)$  に関して、 $1 \geq h(0) \geq h(1) \geq h(2) \geq \dots \geq h(N-1)$  かつ、少なくともどこか 1ヶ所で等号が不成立、 $h(m) \geq 0$ 、 $h(1) > 0$  という条件が成り立つとする。このとき、参照ベクトルの任意の初期配置に対して、有限の長さの入力系列からなる有限の確率測度をもつ集合で、この集合に属するどの入力系列によっても参照ベクトルが整列するようなものが存在する。

**証明** 参照ベクトルを  $w_i \in I = [a, b]$ ,  $i = 1, \dots, N$  で表わす。まず、基本的な事項からみていく。

**補題 1** 参照ベクトルが整列している状態は、学習則によって乱されない。

**証明**  $1 \geq h(0) \geq h(1) \geq h(2) \geq \dots \geq h(N-1)$  であれば成り立つことは明らかである。(証明おわり)

**定義 1**  $1 \leq i < j \leq N$  とし、 $k < i$  または  $k > j$  であるような  $k$  のそれぞれに対して「 $w_k < \min_{i \leq l \leq j} w_l$  または  $w_k > \max_{i \leq l \leq j} w_l$ 」が成り立つとき、 $\{w_i, \dots, w_j\}$  を  $(i, j)$ -クラスタという。

**定義 2** 全要素が整列している  $(i, j)$ -クラスタを  $(i, j)$ -ブロックという。

**定義 3**  $w_i < \dots < w_j$  あるいは  $w_i > \dots > w_j$  のとき、これを  $(i, j)$ -チェーンという。

**補題 2**  $(i, j)$ -チェーンは、入力信号  $x \in V_k$ ,  $i \leq k \leq j$  によって乱されない。

以下の議論では、 $h(D) \geq h(D+1)$ ,  $2D < N$  とする。

補題 3  $|j| \leq D < |k|$  とする. 任意の  $i, 1 \leq i \leq N$  に対して, 入力信号  $x \in V_i$  を有限回とって

$$|w_i - w_{i+j}| < |w_i - w_{i+k}| \quad (\text{C.1})$$

とできる.

系 3.1 任意の  $i, 1 \leq i \leq N$  に対して, 入力信号  $x \in V_i$  を有限回とって  $(i-D, i+D)$ -クラスタをつくることができる.

定理の証明に移る. 初期状態で  $w_1 < w_{D+2}$  であるとしても一般性を失わない.

- (1) 入力信号  $x \in V_{D+2}$  を有限回とることで,  $(2, 2D+2)$ -クラスタをつくることができる (補題 3 の系 3.1). この時点で  $w_1 < \min_{2 \leq j \leq 2D+2} w_j$  だから, とくに  $w_1 < w_2, w_1 < w_{D+3}$  である.  $(1, 2)$ -チェーンができています.
- (2) 入力信号  $x \in V_1$  を有限回とって,  $(1, D+1)$ -クラスタをつくることができる.  $w_1 < w_{D+3}$  だったから, この時点では  $w_2 < w_{D+3}$  が成り立っている.  $(1, 2)$ -チェーンは保存されている (補題 2).
- (3-1)  $(1, k)$ -チェーンができていて,  $w_k < w_{k+D+1}$  が成り立っているとしよう ( $k \leq N - D - 2$  とする).
- (3-2) 入力信号  $x \in V_{k+D+1}$  を有限回とって,  $(k+1, k+2D+1)$ -クラスタをつくることができる.  $w_k < w_{k+D+1}$  だったから, この時点で  $w_k < w_{k+1}, w_k < w_{k+D+2}$  が成り立っている. したがって  $(1, k+1)$ -チェーンができる.
- (3-3) 入力信号  $x \in V_k$  を有限回とって,  $(k-D, k+D)$ -クラスタをつくることができる.  $w_k < w_{k+D+2}$  だったから, この時点では  $w_{k+1} < w_{k+D+2}$  が成り立っている.  $(1, k+1)$ -チェーンは保存されている (補題 2).
- (4) (3-1) から (3-3) までを  $k = N - D - 1$  となるまで繰り返して,  $(1, k+1)$ -チェーンができていて残りの参照ベクトルはすべて  $w_{k+1}$  より大きくなるようにできる.

(5)  $(1, l)$ -チェーンができていて ( $l \geq N - D - 1$  とする),  $l < i \leq N$  に対して  $w_i > w_l$  が成り立っているとす。このとき,  $(1, l)$ -ブロックができてい。条件  $2D < N$  より  $l \geq D$  がいえるから, 入力信号  $x \in V_{l-D+1}$  を有限回とって,  $(l - 2D + 1, l + 1)$ -クラスタをつくることができる。  $w_l < w_{l+1}$  だったから, これによって  $(1, l + 1)$ -ブロックができたことになる。さらに  $l + 1 < i < N$  に対して  $w_i > w_{l+1}$  が成り立っている。

(6) (5) を繰り返すことによって,  $(1, N)$ -ブロックができる。このブロックはそれ以降の入力信号によって乱されない。

すなわち,  $N$  個の参照ベクトルが整列するような入力系列の存在が示されたことになる。

上で述べた手続きをみたす入力系列は有限の長さである。さらに, このような入力系列全体の確率測度は明らかに 0 より大きい。(証明おわり)