

ニューラルネットワーク (ISBN 4-254-11612-8) 自習 ノート

目次

1	ニューラルネットワークとは何か	3
1.1	生物に学ぶ	3
1.1.1	脳の神経細胞	3
1.1.2	神経細胞の構造と機能	3
1.2	神経細胞のモデル	5
1.3	シナプスの可塑性	5
1.4	ニューラルネットワークの分類	5
1.4.1	階層型ニューラルネットワーク	5
1.4.2	相互結合型ニューラルネットワーク	5
1.5	ニューラルネットワークの特徴	5
1.5.1	並列分散処理	5
1.5.2	学習と自己組織化	5
2	階層型ニューラルネットワーク	5
2.1	パーセプトロン	5
2.1.1	単純パーセプトロン	5
2.1.2	単純パーセプトロンの学習	7
2.2	バックプロパゲーション	14
2.2.1	一般化デルタ則	14
2.2.2	誤差逆伝搬法	16
2.2.3	応用例	19
2.2.4	ニューラルネットワークの構造とパラメータの与え方	19

2.2.5	誤差逆伝搬法の改良	19
3	相互結合型ニューラルネットワーク	19
3.1	相互結合型ニューラルネットワークの形態	19
3.2	連想記憶	19
3.3	ホップフィールドモデル	20
3.3.1	ホップフィールドモデル (Hopfield model)	20
3.3.2	2 値ホップフィールドモデル	21
3.3.3	連想記憶への応用	23
3.3.4	連続値ホップフィールドモデル	24
3.3.5	最適化問題への応用	26
3.3.6	連続値ホップフィールドモデルの改良	28
3.4	ボルツマンマシン	28
3.4.1	ボルツマンマシンの動作	28
3.4.2	ボルツマンマシンの学習	28
3.4.3	ボルツマンマシンの特徴	28
4	競合学習型ニューラルネットワーク	28
4.1	認識機構の自己形成	28
4.2	生体のトポロジカルマッピングのモデル	28
4.3	コホーネンのモデル	28
4.3.1	予備実験	28
4.3.2	特徴抽出細胞の形成	28
4.3.3	コホーネンの学習則	28
4.3.4	コホーネンの自己組織化特徴マップのアルゴリズムとシミュレーション	28
4.3.5	応用例	28
5	ニューラルネットワーク研究の意義	28
5.1	特徴を生かす	28
5.1.1	研究の歴史	28
5.1.2	生物内のニューラルネットワークと人工ニューラルネットワーク	28
5.1.3	シナプスの可塑性と脳・神経系の可塑性	28
5.1.4	教師あり学習と教師なし学習	28
5.1.5	ニューロンコンピュータ	28
5.1.6	融合化技術	28

5.2	応用	28
5.2.1	応用されてきた分野	28
5.2.2	事例の完備性と適用有効範囲	28
5.2.3	ブラックボックスモデルの利用環境への適合性	28
5.3	脳科学への貢献	28

1 ニューラルネットワークとは何か

1.1 生物に学ぶ

1.1.1 脳の神経細胞

神経細胞には大きくわけて

- 錐体細胞 (pyramidal cell)
- プルキンエ細胞 (purkinje cell)
- 星状膠細胞 (astrocyte)

の3つがある。この3つのうち錐体細胞は脳の70から85%を占める。神経細胞を三角で表すことがあるのは、この錐体細胞が三角形に見えるからである。

1.1.2 神経細胞の構造と機能

神経細胞は大きく

細胞体 (soma) 核 (nucleus) を細胞膜 (cell membrane) が包んだもの

樹状突起 (dendrite) 樹状突起は細胞体からつきだしているもの

軸索 (axon) 他の神経細胞の細胞体もしくは樹状突起へ繋がっているもの

によって構成される。軸索と他の神経細胞との結合部分をシナプス (synapse) という。

体液内と細胞内にはイオンが存在している。体内のイオンのうち神経細胞の働きに大きく影響を与えるのはナトリウムイオンとカリウムイオンである。通常、体液内にはナトリウムイオンが多く存在し、細胞内にはカリウムイオンが多く存在する。また細胞膜にはイオンチャネルが存在し、イオンチャネルが開いていると特定のイオンがイオンチャネルを通過することができる。細胞内と体液内には電位差があり、体液内に対する細胞内の電位を膜電位 (membrane potential) という。

神経細胞の細胞膜に存在するカリウムイオンチャネルは通常開いており、カリウムイオンが常に細胞内と体液内を行き来することができる。他に、電位依存性のナトリウムイオン

チャンネルや電位依存性のカリウムイオンチャンネルが存在し、どちらも通常は閉じている。また細胞は DNA によってタンパク質を常に製造しており、タンパク質は負の電荷を帯びている。よって体液内のカリウムイオン (1 荷の正イオン) が体液内と細胞内の電位差によって細胞内に入ってくる。しかし、一定数のカリウムイオンが細胞内に入ると濃度の差によりカリウムイオンは細胞内から体液内へと移動する。通常時はこのカリウムイオンの出入りが安定した状態になる。この状態の膜電位を静止電位 (resting potential) といい一般に -70mV ほどである。

膜電位がある一定の値 (およそ -50mV から -40mV) まで上がると、細胞膜上の電位依存性のナトリウムイオンチャンネルが開く。すると電位差によりナトリウムイオン (1 荷の正イオン) が体液内から細胞内に入ってくる。これにより膜電位は大きく上昇し、膜電位が約 40mV ほどになると電位依存性のナトリウムイオンチャンネルは閉じる。それとほぼ同時に電位依存性のカリウムイオンチャンネルが開く。このとき、膜電位が正であることと細胞内にカリウムイオンが多く存在することから、カリウムイオンは細胞内から体液内へと移動する。これにより膜電位は大きく下がる。膜電位が約 -70mV 程度 (もうちょっと小さく?) になると電位依存性のカリウムイオンチャンネルは閉じる。これ以降はカリウムイオンチャンネルのみが開いている状態のため通常の状態に戻り膜電位は静止電位に戻る。膜電位が一定の値を越えてからのこの一連の流れが発生することを発火 (fire) するもしくは興奮 (excite) するという。また、神経細胞が発火する膜電位の値のことをしきい値 (threshold) といい、発火したときの膜電位の値を活動電位 (action potential) という。

発火している時間はだいたい同じであり、1 度しきい値を越えれば自動的に膜電位が増減するため同じような膜電位の増減のしかたをする。また、発火している最中に重ねて発火するようなことは起きないため発火すればどの神経細胞も同じような膜電位の増減のしかたをする。この膜電位の増減は細胞膜上をつたわる。この膜電位の上昇が軸索をとおりシナプスへ伝わることで、シナプスが別の神経細胞へ信号を伝える。シナプスが他の神経細胞へ信号を伝える際、シナプスは伝達物質を用いるが、これには興奮性の物質と抑制性の物質がある。どちらの物質を出すかはシナプスによって決まっている。神経細胞は他の神経細胞とのシナプスから伝達物質を受け取ると膜電位が上昇する。

神経細胞を信号処理器と捉えると、軸索 (もしくはシナプス) は出力、細胞膜や樹状突起は入力とみることができる。複数の他神経細胞から伝達物質を受取り膜電位がしきい値を越えると、発火しそれを他の神経細胞へ伝えるという過程は非線形なふるまいであり、これが脳の高性能な処理能力の所以であるとされる。

1.2 神経細胞のモデル

1.3 シナプスの可塑性

1.4 ニューラルネットワークの分類

1.4.1 階層型ニューラルネットワーク

1.4.2 相互結合型ニューラルネットワーク

1.5 ニューラルネットワークの特徴

1.5.1 並列分散処理

1.5.2 学習と自己組織化

2 階層型ニューラルネットワーク

階層型ニューラルネットワークは心理学者ローゼンブラットによって提案されたパーセプトロンから発展したものである。パーセプトロンはパターンを学習・識別することができるニューラルネットワークであり、形式ニューロンとシナプスの可塑性を用いている。また、小脳においてパーセプトロンと類似した機能を有する部分があると指摘されている。本章ではパーセプトロンから発展した階層型ニューラルネットワークについて述べる。ニューラルネットワークによる応用のうちで最も多いのが、階層型ニューラルネットワークのバックプロパゲーションである。

2.1 パーセプトロン

2.1.1 単純パーセプトロン

パーセプトロンは3つの層からなる階層型ニューラルネットワークである。パーセプトロンを構成する3つの層はそれぞれ感覚ユニット (sensory unit)、連合ユニット (associate unit)、反応ユニット (response unit) と呼ばれる。パーセプトロンの基本形である単純パーセプトロンは感覚ユニットから連合ユニット、連合ユニットから反応ユニットへと一方向に繋がっている。また反応ユニットは1つのニューロンによって構成され、連合ユニットの全てのニューロンと接続される。

感覚ユニットの各ニューロンは入力を $x \in \{0,1\}$ としたとき出力値 z は

$$z = x$$

と表わされる。連合ユニットと反応ユニットのニューロンはマッカロとピッツの形式ニューロンであり、入力 (神経細胞が受け取るシナプス前神経細胞の出力) を $x_k \in \{0,1\}$ ($k =$

$1, 2, \dots, n$), $w_k \in \mathbb{R}$ をシナプス結合荷重、 y を膜電位の変化量としたとき、出力値 z は

$$y = \sum_{k=1}^n w_k x_k \quad (2.1)$$

$$f(u) = \begin{cases} 1, & u > 0 \\ 0, & u \leq 0 \end{cases} \quad (2.2)$$

$$z = f(y - h) \quad (2.3)$$

という計算式で表わされる。ここで感覚ユニットのニューロン数を m とし、 i 番目 ($i = 1, 2, \dots, m$) のニューロンの出力を o_1^I, \dots, o_m^I 、連合ユニットのニューロン数を n として感覚ユニットの i 番目のニューロンと連合ユニットの j 番目 ($j = 1, 2, \dots, n$) のユニットとのシナプス結合荷重を $w_{i,j}^{I,M}$ とすると連合ユニットの j 番目のニューロンへの入力

$$\sum_{i=1}^m w_{i,j}^{I,M} o_i^I$$

と表わすことができる。この式は式 (2.1) に対応し連合ユニットの各ニューロンの出力値は式 (2.2)、式 (2.3) によって計算されるため、 θ_j^M を連合ユニットの j 番目のニューロンのしきい値とすると連合ユニットのニューロンの出力値 o_1^M, \dots, o_n^M は

$$o_j^M = \begin{cases} 1, & \sum_{i=1}^m w_{i,j}^{I,M} o_i^I - \theta_j^M > 0 \\ 0, & \sum_{i=1}^m w_{i,j}^{I,M} o_i^I - \theta_j^M \leq 0 \end{cases}$$

となる。同様に連合ユニットの j 番目のニューロンと反応ユニットのニューロンとのシナプス結合荷重を $w_j^{M,O}$ 、反応ユニットのニューロンのしきい値を θ^O とすれば、反応ユニットのニューロンの出力値 o^O は

$$o^O = \begin{cases} 1, & \sum_{j=1}^n w_j^{M,O} o_j^M - \theta^O > 0 \\ 0, & \sum_{j=1}^n w_j^{M,O} o_j^M - \theta^O \leq 0 \end{cases} \quad (2.4)$$

となる。また、この単純パーセプトロンにおいて学習するパラメータは $w_j^{M,O}$ ($j = 1, 2, \dots, n$) と θ^O のみであり、 $w_{i,j}^{I,M}$ ($i = 1, 2, \dots, m, j = 1, 2, \dots, n$) と θ_j^M は固定して学習を行う。

ここで $w_{n+1}^{M,O} = -\theta^O$ 、 $o_{n+1}^M = 1$ とすれば

$$\begin{aligned} \sum_{j=1}^n w_j^{M,O} o_j^M - \theta^O &= \sum_{j=1}^n w_j^{M,O} o_j^M + w_{n+1}^{M,O} o_{n+1}^M \\ &= \sum_{j=1}^{n+1} w_j^{M,O} o_j^M \end{aligned}$$

と書きなおすことができる。また、新たに $\mathbf{w}^{M,O} \in \mathbb{R}^{n+1}$, $\mathbf{o}^M \in \{0,1\}^{n+1}$ を

$$\begin{aligned}\mathbf{w}^{M,O} &= (w_1^{M,O}, w_2^{M,O}, \dots, w_n^{M,O}, w_{n+1}^{M,O}) = (w_1^{M,O}, w_2^{M,O}, \dots, w_n^{M,O}, -\theta^O) \\ \mathbf{o}^M &= (o_1^M, o_2^M, \dots, o_n^M, o_{n+1}^M) = (o_1^M, o_2^M, \dots, o_n^M, 1)\end{aligned}$$

とするとベクトル同士の内積によって

$$\sum_{j=1}^n w_j^{M,O} o_j^M - \theta^O = \sum_{j=1}^{n+1} w_j^{M,O} o_j^M = \mathbf{w}^{M,O} \cdot \mathbf{o}^M \quad (2.5)$$

が成立する。

2.1.2 単純パーセプトロンの学習

単純パーセプトロンの学習法のうち、代表的なものである誤り訂正法 (learning by error-correction) について説明する。誤り訂正法は、複数の入力と教師信号 (入力を与えた際に求められる出力値) の組を用いて学習を行う。おおまかな手順は以下の通りである。

1. パーセプトロンに入力を与えて出力値を計算する
2. 得られた出力値と教師信号を比較してシナプス結合荷重としきい値によるベクトル $\mathbf{w}^{M,O}$ を修正する
3. 修正がなくなるまで 1 へ戻る

シナプス結合荷重としきい値によるベクトル $\mathbf{w}^{M,O}$ を修正した $\mathbf{w}'^{M,O}$ は

$$\Delta \mathbf{w}^{M,O} = \eta (t^O - o^O) \mathbf{o}^M$$

によって計算される $\Delta \mathbf{w}^{M,O}$ を用いて

$$\mathbf{w}'^{M,O} = \mathbf{w}^{M,O} + \Delta \mathbf{w}^{M,O} \quad (2.6)$$

と表わされる。ここで $t^O \in \{0,1\}$ は入力に対する教師信号、 $\eta \in (0,1]$ は学習率 (learning rate) である。学習率は学習に大きく影響を与える定数であり、実際に学習する際には学習率 η は適切に決める必要がある。

t^O, o^O 共に 0,1 のどちらかであるため $t^O - o^O$ がとりうる値は

$$t^O - o^O = \begin{cases} 1, & t^O = 1, o^O = 0 \\ -1, & t^O = 0, o^O = 1 \\ 0, & t^O = o^O \end{cases}$$

であり $\Delta \mathbf{w}^{M,O}$ は

$$\Delta \mathbf{w}^{M,O} = \begin{cases} \eta \mathbf{o}^M, & t^O = 1, o^O = 0 \\ -\eta \mathbf{o}^M, & t^O = 0, o^O = 1 \\ 0, & t^O = o^O \end{cases}$$

と書くことができる。また、教師信号 t^O は入力に対して o^O がとるべき値であるため、式 (2.6) はパーセプトロンが入力に対して誤った値を出力したときにシナプス結合荷重としきい値を修正するという式である。さらに修正後の $\mathbf{w}^{M,O}$ を用いて同じ入力に対する式 (2.5) の値を計算すると

$$\begin{aligned} \mathbf{w}^{M,O} \cdot \mathbf{o}^O &= (\mathbf{w}^{M,O} + \Delta \mathbf{w}^{M,O}) \cdot \mathbf{o}^M \\ &= \mathbf{w}^{M,O} \cdot \mathbf{o}^O + \Delta \mathbf{w}^{M,O} \cdot \mathbf{o}^M \\ &= \mathbf{w}^{M,O} \cdot \mathbf{o}^O + \left[\eta (t^O - o^O) \mathbf{o}^M \right] \cdot \mathbf{o}^M \\ &= \mathbf{w}^{M,O} \cdot \mathbf{o}^O + \eta (t^O - o^O) \mathbf{o}^M \cdot \mathbf{o}^M \\ &= \mathbf{w}^{M,O} \cdot \mathbf{o}^O + \eta (t^O - o^O) \|\mathbf{o}^M\|^2 \end{aligned}$$

となり $t^O - o^O$ がとりうる値から、 $t^O = 0, o^O = 1$ のときは式 (2.5) の値は $\eta \|\mathbf{o}^M\|^2$ だけ減ることで o^O が 0 になりやすくなり、 $t^O = 1, o^O = 0$ のときは式 (2.5) の値は $\eta \|\mathbf{o}^M\|^2$ だけ増えることで o^O が 1 になりやすくなり、それ以外の場合は式 (2.5) の値と o^O の値は変化しないことがわかる。

学習をつづけることを時間の経過とらえることで式 (2.5) はヘブの強化則ということができ、式 (2.5) はシナプス結合荷重を変化させているためシナプスの可塑性を利用しているともいうことができる。

式 (2.6) によるシナプス結合荷重の修正を複数の入力と教師信号の組に対して繰り返し行なうことで学習を進める。誤り訂正法による単純パーセプトロンの学習の詳細な手順は以下の通りである。

- Step1 複数の入力パターン $\mathbf{i}_p^I = (i_{p,1}^I, i_{p,2}^I, \dots, i_{p,m}^I) \in \mathbb{R}^m$ ($p = 1, 2, \dots, P$) と、 p 番目の入力パターンと対応する教師信号 $t_p^O \in \{0, 1\}$ を用意する。
- Step2 シナプス結合荷重 $\mathbf{w}^{M,O} = (w_1^{M,O}, w_2^{M,O}, \dots, w_{n+1}^{M,O})$ の初期値を乱数によって小さい値に設定する。そして、学習率 $\eta \in (0, 1]$ を与える。
- Step3 ある $p \in \{1, 2, \dots, P\}$ に対して、 p 個目の入力パターン \mathbf{i}_p^I に対する連合ユニットの出力値 $o_{p,j}^M$ ($j = 1, 2, \dots, n$) を

$$o_{p,j}^M = \begin{cases} 1, & \sum_{i=1}^m w_{i,j}^{I,M} i_{p,i}^I - \theta_j^M > 0 \\ 0, & \sum_{i=1}^m w_{i,j}^{I,M} i_{p,i}^I - \theta_j^M \leq 0 \end{cases}$$

によって計算する。

Step4 得られた連合ユニットのニューロンの出力 $o_{p,j}^{M,O}$ ($j = 1, 2, \dots, n$) によって表わされるベクトル $\mathbf{o}_p^M = (o_{p,1}^M, o_{p,2}^M, \dots, o_{p,n}^M, 1)$ とシナプス結合荷重から反応ユニットのニューロンの出力値 o_p^O を

$$o_p^O = \begin{cases} 1, & \mathbf{w}^{M,O} \cdot \mathbf{o}_p^M > 0 \\ 0, & \mathbf{w}^{M,O} \cdot \mathbf{o}_p^M \leq 0 \end{cases}$$

によって計算する。

Step5 反応ユニットのニューロンの出力値の計算に用いた入力パターンベクトル \mathbf{t}_p^I と対応する教師信号 t_p^O と得られた反応ユニットの出力値 o_p^O を用いて計算される $\Delta \mathbf{w}^{M,O}$ を

$$\Delta \mathbf{w}^{M,O} = \eta (t_p^O - o_p^O) \mathbf{o}_p^M$$

によって計算する。そして、シナプス結合荷重 $\mathbf{w}^{M,O}$ に $\Delta \mathbf{w}^{M,O}$ を加算することでシナプス結合荷重を修正する。

Step6 任意の $p \in \{1, 2, \dots, P\}$ に対して、入力パターンベクトル \mathbf{i}_p^I から計算される反応ユニットの出力値 o_p^O と t_p^O が等しいならば終了する。そうでない場合は [Step3](#) から [Step5](#) までを繰り返す。

実際に学習を行う際には [Step6](#) において、シナプス結合荷重の変化 $\Delta \mathbf{w}^{M,O}$ が一定の値より小さくなったら終了としたり、もしくは繰り返し回数に上限を設けて繰り返し回数が上限を越えたときに終了とするなどの条件も用いる。[Step2](#) において学習率 η の値は大きいほど学習が早く進むがシナプス結合荷重や、教師信号と計算された反応ユニットの出力値との差が発散、振動しやすくなる。そのため η には適切な値を用いる必要がある。適切な η の値は問題の対象によって変化する。

単純パーセプトロンは入力パターンを 2 つの集合に分類することができる機械であり、分類する事を認識という。単純パーセプトロンにおける認識とは、入力パターンに対して反応ユニットのニューロンの出力値を計算することであり、出力値が 0 か 1 かによって入力パターンが 2 つの集合に分類される。式 (2.4) から反応ユニットのニューロンの出力値は式 (2.5) の値が正か負かによって決まることがわかる。式 (2.5) が

$$\sum_{j=1}^n w_j^{M,O} o_j^M - \theta^O = \sum_{j=1}^{n+1} w_j^{M,O} o_j^M = \mathbf{w}^{M,O} \cdot \mathbf{o}^M = 0$$

という $n+1$ 次元の超平面を表わしていると考え、認識することは入力パターンと対応する \mathbf{o}^M を超平面によって 2 つの集合にわけると言い換えることができる。認識するこ

とは、入力パターン $\mathbf{i}^I \in \{0,1\}^m$ を座標変換によって $\mathbf{o}^M \in \{0,1\}^{n+1}$ へと変換し \mathbf{o}^M を n 次元の超平面によって 2 つの集合のどちらかに分類わけすることと言い換えることができる。

また、単純パーセプトロンにおいて学習するパラメータはシナプス結合荷重 $\mathbf{w}^{M,O}$ のみであるため、入力パターン \mathbf{i}^I から \mathbf{o}^M への座標変換のパラメータは変化せずに n 次元の超平面の係数のみが変わることとなる。したがって、単純パーセプトロンの学習は入力パターンを適切に識別可能な n 次元超平面を求めることと同値となる。問題によってはどのような超平面を用いても正しく識別することができないものが存在し、逆に正しく識別することができる超平面が存在するとき線形分離可能という。

定義 1：線形分離可能 (linear separable) $F^+, F^- \in \{0,1\}^n$ ($n \in \mathbb{N}$) としたとき任意の $\mathbf{o}^+ \in F^+, \mathbf{o}^- \in F^-$ に対して

$$\begin{aligned}\mathbf{w} \cdot \mathbf{o}^+ &> 0 \\ \mathbf{w} \cdot \mathbf{o}^- &\leq 0\end{aligned}$$

となる $\mathbf{w} \in \mathbb{R}^n$ が存在することを F^+ と F^- は線形分離可能であるという。

パーセプトロンでは線形分離可能でない問題に対して適切な識別を行なうことはできない。したがってパーセプトロンが扱うのは線形分離可能な問題である。また線形分離可能な問題に対してであれば誤り訂正法を用いて必ず適切なパラメータを学習することができることが示される。

定理 1：パーセプトロンの収束定理 (perceptron convergence theorem)

定理 1 の証明. 正しい識別を与えるシナプス結合荷重のうち任意の $\mathbf{o} \in F^+ \cup F^-$ に対して

$$\mathbf{o} \cdot \mathbf{w}^* \neq 0$$

となる $\mathbf{w}^* \in \mathbb{R}^{n+1}, \mathbf{w}^* \neq \mathbf{0}$ をひとつ選ぶ。実数空間が連続であることから \mathbf{w}^* は必ず存在する。

また、シナプス結合荷重の初期値を \mathbf{w}_0 、 t 回修正を繰り返したシナプス結合荷重を \mathbf{w}_t とすると式 (2.6) から、ある $\mathbf{o}^M \in \{0,1\}^{n+1}$ が存在して

$$\mathbf{w}_{t+1} = \mathbf{w}_t + \eta (t^O - o^O) \mathbf{o}^M \quad (2.7)$$

と書くことができる。ここで t^O は \mathbf{o}^M と対応する教師信号であり、 o^O は \mathbf{o}^M を連合ユニットのニューロンの出力値からなるベクトル、 \mathbf{w}_t を連合ユニットと反応ユニット間のシナプス結合荷重として計算された反応ユニットの出力値である。

式 (2.7) の両辺と \mathbf{w}^* の内積をとり、変形すると

$$\begin{aligned}\mathbf{w}^* \cdot \mathbf{w}_{t+1} &= \mathbf{w}^* \cdot \left[\mathbf{w}_t + \eta \left(t^O - o^O \right) \mathbf{o}^M \right] \\ &= \mathbf{w}^* \cdot \mathbf{w}_t + \eta \left(t^O - o^O \right) \mathbf{w}^* \cdot \mathbf{o}^M\end{aligned}\quad (2.8)$$

となる。式 (2.8) は修正が発生したときの式であるため $t^O \neq o^O$ であり $(\mathbf{w}^* \cdot \mathbf{o}^M) (\mathbf{w}_t \cdot \mathbf{o}^M) \leq 0$ である。したがって式 (2.4) から

$$\begin{aligned}o^O &= \begin{cases} 0, & \mathbf{w}^* \cdot \mathbf{o}^M > 0 \\ 1, & \mathbf{w}^* \cdot \mathbf{o}^M \leq 0 \end{cases} \\ t^O &= \begin{cases} 1, & \mathbf{w}^* \cdot \mathbf{o}^M > 0 \\ 0, & \mathbf{w}^* \cdot \mathbf{o}^M \leq 0 \end{cases}\end{aligned}$$

であり、 $\eta > 0$ であることと \mathbf{w}^* の定義から

$$\eta \left(t^O - o^O \right) \mathbf{w}^* \cdot \mathbf{o}^M > 0$$

が成立する。ここで δ を

$$\delta = \min_{\mathbf{o} \in F^+ \cup F^-} \mathbf{w}^* \cdot \mathbf{o}^M$$

とすると δ は非ゼロな実数となり、任意の $\mathbf{o} \in F^+ \cup F^-$ に対して

$$\eta \left(t^O - o^O \right) \mathbf{w}^* \cdot \mathbf{o}^M \geq \eta \delta \quad (2.9)$$

となる。そして式 (2.8) と式 (2.9) から、任意の t ($t = 1, 2, \dots$) に対して

$$\begin{aligned}\mathbf{w}^* \cdot \mathbf{w}_t + \eta \left(t^O - o^O \right) \mathbf{w}^* \cdot \mathbf{o}^M &\geq \mathbf{w}^* \cdot \mathbf{w}_t + \eta \delta \\ \Leftrightarrow \mathbf{w}^* \cdot \mathbf{w}_{t+1} &\geq \mathbf{w}^* \cdot \mathbf{w}_t + \eta \delta\end{aligned}$$

となる。したがって、初期値 \mathbf{w}_0 に n 回修正を行なった \mathbf{w}_n に対して

$$\mathbf{w}^* \cdot \mathbf{w}_n \geq \mathbf{w}^* \cdot \mathbf{w}_0 + n\eta\delta \quad (2.10)$$

が成立する。

また式 (2.6) から $\|\mathbf{w}_t\|^2$ は \mathbf{w}_t によって

$$\begin{aligned}\|\mathbf{w}_{t+1}\|^2 &= \left\| \mathbf{w}_t + \eta \left(t^O - o^O \right) \mathbf{o}^M \right\|^2 \\ &= \|\mathbf{w}_t\|^2 + 2\eta \left(t^O - o^O \right) \mathbf{w}_t \cdot \mathbf{o}^M + \eta^2 \left(t^O - o^O \right)^2 \|\mathbf{o}^M\|^2\end{aligned}\quad (2.11)$$

と表わすことができ、式 (2.8) と同様に式 (2.11) は修正が発生したときの式であるため $t^O \neq o^O$ であり $(\mathbf{w}^* \cdot \mathbf{o}^M)(\mathbf{w}_t \cdot \mathbf{o}^M) \leq 0$ である。したがって式 (2.4) から

$$o^O = \begin{cases} 1, & \mathbf{w}_t \cdot \mathbf{o}^M > 0 \\ 0, & \mathbf{w}_t \cdot \mathbf{o}^M \leq 0 \end{cases} \quad (2.12)$$

$$t^O = \begin{cases} 0, & \mathbf{w}_t \cdot \mathbf{o}^M > 0 \\ 1, & \mathbf{w}_t \cdot \mathbf{o}^M \leq 0 \end{cases} \quad (2.13)$$

であり、 $\eta > 0$ であることから

$$2\eta(t^O - o^O)\mathbf{w}_t \cdot \mathbf{o}^M \leq 0 \quad (2.14)$$

が成立する。ここで M を

$$M = \max_{\mathbf{o} \in F^+ \cup F^-} \|\mathbf{o}^M\|$$

とすると $0 < \eta \leq 1$ と式 (2.12)、式 (2.13)、式 (2.14) から、任意の t ($t = 1, 2, \dots$) と \mathbf{o}^M に対して

$$\begin{aligned} \|\mathbf{w}_t\|^2 + 2\eta(t^O - o^O)\mathbf{w}_t \cdot \mathbf{o}^M + \eta^2(t^O - o^O)^2\|\mathbf{o}^M\|^2 &\leq \|\mathbf{w}_t\|^2 + \eta^2(t^O - o^O)^2\|\mathbf{o}^M\|^2 \\ \|\mathbf{w}_t\|^2 + \eta^2(t^O - o^O)^2\|\mathbf{o}^M\|^2 &\leq \|\mathbf{w}_t\|^2 + \eta^2(t^O - o^O)^2M^2 = \|\mathbf{w}_t\|^2 + \eta^2M^2 \end{aligned}$$

となり、式 (2.11) から

$$\|\mathbf{w}_{t+1}\|^2 \leq \|\mathbf{w}_t\|^2 + \eta^2M^2$$

が得られる。したがって、初期値 \mathbf{w}_0 に n 回修正を行なった \mathbf{w}_n に対して

$$\|\mathbf{w}_n\|^2 \leq \|\mathbf{w}_0\|^2 + n\eta^2M^2 \quad (2.15)$$

が成立する。

さらに \mathbf{w}^* が非ゼロベクトルであることから、式 (2.10) は \mathbf{w}^* と \mathbf{w}_n のなす角を θ とすると

$$\begin{aligned} \mathbf{w}^* \cdot \mathbf{w}_n &\geq \mathbf{w}^* \cdot \mathbf{w}_0 + n\eta\delta \\ \Leftrightarrow \|\mathbf{w}^*\| \|\mathbf{w}_n\| \cos \theta &\geq \mathbf{w}^* \cdot \mathbf{w}_0 + n\eta\delta \\ \Leftrightarrow \|\mathbf{w}^*\| \|\mathbf{w}_n\| &\geq \mathbf{w}^* \cdot \mathbf{w}_0 + n\eta\delta \\ \Leftrightarrow \|\mathbf{w}_n\| &\geq \frac{\mathbf{w}^* \cdot \mathbf{w}_0 + n\eta\delta}{\|\mathbf{w}^*\|} \end{aligned} \quad (2.16)$$

と書きかえることができる。

式 (2.15) と式 (2.16) から

$$\frac{(\mathbf{w}^* \cdot \mathbf{w}_0 + n\eta\delta)^2}{\|\mathbf{w}^*\|^2} \leq \|\mathbf{w}_n\|^2 \leq \|\mathbf{w}_0\|^2 + n\eta^2M^2 \quad (2.17)$$

となる。ここで式 (2.17) を見たす $\|\mathbf{w}_n\|^2$ が存在するためには

$$\frac{(\mathbf{w}^* \cdot \mathbf{w}_0 + n\eta\delta)^2}{\|\mathbf{w}^*\|^2} \leq \|\mathbf{w}_0\|^2 + n\eta^2 M^2 \quad (2.18)$$

が成立している必要があり、式 (2.18) を変形すると

$$\begin{aligned} & (\mathbf{w}^* \cdot \mathbf{w}_0 + n\eta\delta)^2 \leq \|\mathbf{w}^*\|^2 \|\mathbf{w}_0\|^2 + n\eta^2 M^2 \|\mathbf{w}^*\|^2 \\ \Rightarrow & n^2 \eta^2 \delta^2 + n \left[2\eta\delta (\mathbf{w}^* \cdot \mathbf{w}_0) - \eta^2 M^2 \|\mathbf{w}^*\|^2 \right] + (\mathbf{w}^* \cdot \mathbf{w}_0)^2 - \|\mathbf{w}^*\|^2 \|\mathbf{w}_0\|^2 \leq 0 \end{aligned} \quad (2.19)$$

となる。また、式 (2.19) の左辺は下に凸な放物線であるため

$$n^2 \eta^2 \delta^2 + n \left[2\eta\delta (\mathbf{w}^* \cdot \mathbf{w}_0) - \eta^2 M^2 \|\mathbf{w}^*\|^2 \right] + (\mathbf{w}^* \cdot \mathbf{w}_0)^2 - \|\mathbf{w}^*\|^2 \|\mathbf{w}_0\|^2 = 0 \quad (2.20)$$

の解を n_-, n_+ ($n_- \leq n_+$) とすると式 (2.19) が成立することと

$$n_- \leq n \leq n_+$$

が成立することは同値である。

ここで式 (2.20) の判別式 D は

$$\begin{aligned} D &= \left[2\eta\delta (\mathbf{w}^* \cdot \mathbf{w}_0) - \eta^2 M^2 \|\mathbf{w}^*\|^2 \right]^2 - 4\eta^2 \delta^2 \left[(\mathbf{w}^* \cdot \mathbf{w}_0)^2 - \|\mathbf{w}^*\|^2 \|\mathbf{w}_0\|^2 \right] \\ &= \eta^4 M^4 \|\mathbf{w}^*\|^4 - 4\eta^3 \delta M^2 (\mathbf{w}^* \cdot \mathbf{w}_0) \|\mathbf{w}^*\|^2 + 4\eta^2 \delta^2 \|\mathbf{w}^*\|^2 \|\mathbf{w}_0\|^2 \\ &= \left[\eta^4 M^4 \|\mathbf{w}^*\|^2 - 4\eta^3 \delta M^2 (\mathbf{w}^* \cdot \mathbf{w}_0) + 4\eta^2 \delta^2 \|\mathbf{w}_0\|^2 \right] \|\mathbf{w}^*\|^2 \end{aligned} \quad (2.21)$$

となり、式 (2.21) は $\mathbf{w}^* = \frac{2\delta}{\eta M^2} \mathbf{w}_0$ のときのみ 0 となり、それ以外の場合に正となることがわかる。さらに $(\mathbf{w}^* \cdot \mathbf{w}_0)^2 - \|\mathbf{w}^*\|^2 \|\mathbf{w}_0\|^2 \leq 0$ であるため

$$n_- \leq 0 \leq n_+$$

が必ず成立し $\mathbf{w}^* = \frac{2\delta}{\eta M^2} \mathbf{w}_0$ でないとき n は有限の値 n_+ によってバウンドされる。また、 $\mathbf{w}^* = \frac{2\delta}{\eta M^2} \mathbf{w}_0$ であるとき $n_- = n_+ = 0$ であり、 \mathbf{w}^* の正の実数倍ベクトルは正しく分離するため $n = 0$ となる。

以上より、シナプス結合荷重の初期値 \mathbf{w}_0 がゼロベクトルであるときに誤り訂正法によるパーセプトロンの学習が有限回の修正で収束することが示された。□

線形分離可能な問題に対しては単純パーセプトロンは適切に学習することができることが示された。実際に単純パーセプトロンを用いる場合、感覚ユニットと連合ユニット間のシナプス結合荷重は学習せずに固定する。感覚ユニットへ線形分離可能でない入力群が与えられる場合でも、感覚ユニットと連合ユニット間でランダムに定められたシナプス結合荷重によ

る座標変換によって、連合ユニットの出力ベクトルの空間上では線形分離可能な問題へ変換することができるようになる。連合ユニットのニューロン数 n が十分に大きいとき、入力パターン数が $2n$ 以下であるどのような問題であっても線形分離可能となることが示されている。

また、パーセプトロンの学習則である誤り訂正法をデルタ則 (delta rule) ということもある。

2.2 バックプロパゲーション

2.2.1 一般化デルタ則

前節で示したデルタ則 (誤り訂正法) を一般化した、一般化デルタ則を導出する。まず入力から出力を計算する過程を一般化する。

入力層でない層のあるニューロン j は、そのニューロンよりも前の層の任意のニューロンの出力 o_i (i は 1 から前の層のニューロン数までの整数) を受けとる。このとき、ニューロン j は前の層のニューロンとのシナプス結合荷重を前の層のニューロンに乘じ和をとっている。よってニューロン j への入力 u_j を

$$u_j = \sum_i w_{i,j} o_i \quad (2.22)$$

と定める。ここで $w_{i,j}$ は前の層のニューロン i とニューロン j のシナプス結合荷重とした。またしきい値を用いていないが、前節で用いたように前の層に出力値が 1 で固定されたニューロンを仮定することでしきい値を表現することはできる。ニューロン j の出力値 o_j は非線形な実数上の 1 変数関数 f によって

$$o_j = f(u_j) \quad (2.23)$$

として求められる。この関数 f は活性化関数もしくは入出力関数といわれる。ここで f をステップ関数とすることで形式ニューロンの式と等価とすることができる。また、

$$f(x) = \frac{1}{1 + e^{-\epsilon x}}$$

と定義されるシグモイド関数は定義域が実数全体であり、終域が実数上の開区間 $(0,1)$ かつ単調増加連続関数であるという性質からよく用いられる。

次に学習する過程を一般化する。ここで、ある入力パターン p (p は 1 から入力パターン数までの整数) に対してパーセプトロンの出力層のニューロン j が出力すべき値を $t_{p,j}$ 、実際のニューロン j (j は 1 から出力層のニューロン数までの整数) の出力値を $o_{p,j}$ とし、誤差関数 E を二乗誤差によって

$$E = \frac{1}{2} \sum_{p,j} (o_{p,j} - t_{p,j})^2 \quad (2.24)$$

と定める。この誤差関数 E の値が 0 となれば入力パターン全てを正しく識別できていることとなる。したがって、誤差関数 E の値を小さくすることを学習といえることができる。

誤差関数 E は出力層のニューロンの値 $o_{p,j}$ に関する関数である。しかし出力層のニューロンの値 $o_{p,j}$ は出力層と前の層との間のシナプス結合荷重 $w_{i,j}$ の関数であるため、誤差関数 E も $w_{i,j}$ に関する (陰 (implicit) に定義された) 関数であるといえる。したがって、各シナプス結合荷重 $w_{i,j}$ に適切な $\eta > 0$ を用いた

$$\Delta w_{i,j} = -\eta \frac{\partial E}{\partial w_{i,j}} \quad (2.25)$$

を加算することで誤差関数 E の極小値 (もしくは鞍点) へ近づけることができる。これは最急降下法 (steepest descent method もしくは gradient decent method) と呼ばれる。

また、式 (2.22) と式 (2.23) から連鎖律により

$$\frac{\partial E}{\partial w_{i,j}} = \sum_p \frac{\partial E}{\partial o_{p,j}} \frac{\partial o_{p,j}}{\partial u_{p,j}} \frac{\partial u_{p,j}}{\partial w_{i,j}}$$

が得られる。ここで $u_{p,j}$ は入力パターン p に対して計算された出力層のニューロン j の入力値である。さらに式 (2.22) と式 (2.23) と式 (2.24) から

$$\begin{aligned} \frac{\partial E}{\partial o_{p,j}} &= o_{p,j} - t_{p,j} \\ \frac{\partial o_{p,j}}{\partial u_{p,j}} &= \left. \frac{df}{dx} \right|_{x=u_{p,j}} \end{aligned}$$

となり、さらに $o_{p,i}$ を出力層の前の中間層の入力パターン p に対するニューロン i の出力値とすると

$$\frac{\partial u_{p,j}}{\partial w_{i,j}} = o_{p,i}$$

となる。よって式 (2.25) の各因子へ代入することで

$$\Delta w_{i,j} = -\eta \sum_p (o_{p,j} - t_{p,j}) \left. \frac{df}{dx} \right|_{x=u_{p,j}} o_{p,i} \quad (2.26)$$

が得られる。

式 (2.26) は一般化デルタ則と呼ばれる。一般化デルタ則では全ての入力パターンに対する誤差を計算してから、シナプス結合荷重の修正量を決め修正している。しかし、 η が非常に小さい場合は入力パターン毎に修正量を決めて修正する方法でも、全体の修正量はほぼ等しくなる。

2.2.2 誤差逆伝搬法

誤差逆伝搬法はフィードフォワード型の階層型ニューラルネットワークの教師あり学習 (supervised learning) の代表的な方法である。入力パターンに対して計算された出力値と教師信号との誤差が小さくなるように、シナプス結合荷重を修正する。

誤差逆伝播法を用いることで入力層と中間層間での高次への写像も学習することができるため、線形分離可能でない問題に対しても適切なパラメータを学習することができる。

実際の学習は、まず入力パターンに対して各ニューロンの出力値を計算し、その後計算された誤差から逆向きにシナプス結合荷重の修正値を計算する。生物のニューラルネットワークでは信号が逆向きに伝わることはないため、誤差逆伝搬法は実際の生物のモデルとしては適切でない。しかし工学上の応用がききやすいため広く利用されている。

以下の表に誤り訂正法と誤差逆伝搬法の違いをまとめる。

表 1 誤り訂正法と誤差逆伝搬法の違い

	誤り訂正法	誤差逆伝搬法
対象	単純パーセプトロン	フィードフォワード型の階層型ニューラルネットワーク
出力値	2 値	実数上の開区間 (0,1)
入出力関数	ステップ関数	シグモイド関数
学習する値	中間層と出力層の間のシナプス結合荷重と出力層のしきい値	任意のシナプス結合荷重としきい値
終了条件	一定回数繰り返しても収束しない場合	2 乗誤差が一定値以下になること

m 層のフィードフォワード型の階層型ニューラルネットワークを学習する場合を考える。学習に用いる入力パターンは P 個のときを考える。パターン p に対しての k 層目の j 番目のニューロンの出力値を $o_{p,j}^k$ 、入力値を $i_{p,j}^k$ 、 $k-1$ 層目の i 番目のニューロンと k 層目の j 番目のニューロンの間のシナプス結合荷重 $w_{i,j}^{k-1,k}$ 、 k 層目の j 番目のニューロンの入出力関数を f_j^k 、しきい値を θ_j^k 、 k 層目のニューロン数を n_k とすると

$$o_{p,j}^k = f_j^k(i_{p,j}^k) \quad (2.27)$$

$$i_{p,j}^k = \sum_{i=1}^{n_{k-1}} w_{i,j}^{k-1,k} o_{p,i}^{k-1} - \theta_j^k \quad (2.28)$$

という関係式が成立する。ここで k 層目に出力値が常に 1 である $n_k + 1$ 番目のニューロンの存在を仮定することで式 (2.28) は

$$t_{p,j}^k = \sum_{i=1}^{n_{k-1}+1} w_{i,j}^{k-1,k} o_{p,i}^{k-1}$$

と書きかえることができる。

パターン p に対する m 層目の j 番目のニューロンの教師信号を $t_{p,i}^{m,k}$ し、誤差関数 E を

$$E = \sum_p E_p$$

$$E_p = \frac{1}{2} \sum_{i=1}^{n_m} \left(t_{p,i}^m - o_{p,i}^m \right)^2$$

と定義する。ここで E_p はパターン p に対する誤差ということができる。

最急降下法によってシナプス結合荷重に

$$\Delta_p w_{i,j}^{k-1,k} = -\eta \frac{\partial E_p}{\partial w_{i,j}^{k-1}} \quad (2.29)$$

を加算することで修正する。 η は半開区間 $(0, 1]$ 上の実数である。

式 (2.29) の右辺の偏微分は連鎖律によって

$$\frac{\partial E_p}{\partial w_{i,j}^{k-1}} = \frac{\partial E_p}{\partial i_{p,j}^k} \frac{\partial i_{p,j}^k}{\partial w_{i,j}^{k-1}}$$

と書くことができ、

$$\frac{\partial i_{p,j}^k}{\partial w_{i,j}^{k-1}} = o_{p,i}^{k-1}$$

であるため

$$\frac{\partial E_p}{\partial w_{i,j}^{k-1}} = \frac{\partial E_p}{\partial i_{p,j}^k} o_{p,i}^{k-1}$$

となり、

$$\delta_{p,j}^k = -\frac{\partial E_p}{\partial i_{p,j}^k}$$

と置くことで

$$-\frac{\partial E_p}{\partial w_{i,j}^{k-1}} = \delta_{p,j}^k o_{p,i}^{k-1}$$

となる。最急降下法を用いるために非ゼロな 1 以下の実数定数を学習率 η として、 $w_{i,j}^{k-1,k}$ のパターン p に対する修正量 $\Delta_p w_{i,j}^{k-1,k}$ は

$$\Delta_p w_{i,j}^{k-1,k} = \eta \delta_{p,j}^k o_{p,i}^{k-1}$$

と表わすことができる。

また、 $\delta_{p,j}^k$ は連鎖律と式 (2.27) によって

$$\begin{aligned} \delta_{p,j}^k &= -\frac{\partial E_p}{\partial i_{p,j}^k} \\ &= -\frac{\partial E_p}{\partial o_{p,j}^k} \frac{\partial o_{p,j}^k}{\partial i_{p,j}^k} \\ &= -\frac{\partial E_p}{\partial o_{p,j}^k} \left. \frac{df_j^k(x)}{di_{p,j}^k} \right|_{x=i_{p,j}^k} \end{aligned}$$

と書くことができる。ここで

$$\frac{\partial E_p}{\partial o_{p,j}^k} = \begin{cases} o_{p,j}^m - t_{p,j}^m & k = m \\ \sum_{l=1}^{n_k+1} \left(\frac{\partial E_p}{\partial i_{p,l}^{k+1}} \frac{\partial i_{p,j}^{k+1}}{\partial o_{p,j}^k} \right) = -\sum_{l=1}^{n_k+1} \left(\delta_{p,l}^{k+1} w_{j,l}^{k,k+1} \right) & \text{otherwise} \end{cases}$$

となり、まとめると層数 m の階層型ニューラルネットワークの学習は

$$\begin{aligned} \Delta_p w_{i,j}^{k-1,k} &= \eta \delta_{p,j}^k o_{p,i}^{k-1} \\ i &= 1, 2, \dots, n_k + 1, \quad j = 1, 2, \dots, n_k, \quad k = 2, 3, \dots, m \end{aligned}$$

$$\delta_{p,j}^k = -\frac{\partial E_p}{\partial o_{p,j}^k} \left. \frac{df_j^k(x)}{di_{p,j}^k} \right|_{x=i_{p,j}^k}$$

$$\frac{\partial E_p}{\partial o_{p,j}^k} = \begin{cases} o_{p,j}^m - t_{p,j}^m & k = m \\ \sum_{l=1}^{n_k+1} \left(\frac{\partial E_p}{\partial i_{p,l}^{k+1}} \frac{\partial i_{p,j}^{k+1}}{\partial o_{p,j}^k} \right) = -\sum_{l=1}^{n_k+1} \left(\delta_{p,l}^{k+1} w_{j,l}^{k,k+1} \right) & \text{otherwise} \end{cases}$$

として表わされる $\Delta_p w_{i,j}^{k-1,k}$ を $w_{i,j}^{k-1,k}$ に各パターンに対して加算することで行われる。

したがって、誤差逆伝搬法において入出力関数は C^1 級関数である必要がある。誤差逆伝搬法で一般に用いられるシグモイド関数は

$$f(x) = \frac{1}{1 + e^{-\epsilon x}}$$

と定められる実数全体から実数区間 $(0, 1)$ への単調増加連続関数である。また、シグモイド関数の 1 階導関数は

$$\frac{df(x)}{dx} = \epsilon f(x)(1 - f(x))$$

である。

学習の際に修正するとき、各パターン毎に修正する方法と全パターンの修正量をまとめて後から修正する方法の 2 つがある。前者を逐次修正法 (またはオンライン学習)、後者を一括修正法 (またはバッチ学習もしくはフルバッチ学習) という。また、逐次修正法と一括修正法を合わせて、いくつかのパターン分の修正をまとめて行なうミニバッチ学習という方法も用いられる。厳密に逐次修正法は全パターンに対する誤差の極小値を得るわけではないが学習率 η を小さくすることで全パターンに対する誤差の極小値を得ることができる。

2.2.3 応用例

誤差逆伝搬法は広く応用に使われている。応用されている事例をいくつか示す。

英語の発音学習システム a

2.2.4 ニューラルネットワークの構造とパラメータの与え方

2.2.5 誤差逆伝搬法の改良

3 相互結合型ニューラルネットワーク

3.1 相互結合型ニューラルネットワークの形態

相互結合型ニューラルネットワークは任意のニューロン同士が結合しているネットワークである。階層型ニューラルネットワークは一般に結合の向きで各ニューロンに順序関係が定まるのに対して、相互結合型ニューラルネットワークにおいては順序関係が定まらない。また、階層型ニューラルネットワークは一般に時系列を扱わないのに対して、相互結合型ニューラルネットワークは時刻の経過によって出力を計算する。

相互結合型ニューラルネットワークには外界から入力を受け取るユニット I 、外界へ出力を出すユニット O 、それ以外のユニット H が存在する。また必ずしも $I \cup O = \phi$ であるとは限らない。

3.2 連想記憶

人間における記憶の検索は連想によって行なわれているとされる。この連想による検索を再現するニューラルネットワークが相互結合型ニューラルネットワークである。ここで、

ニューラルネットワークにおける連想記憶の定義を示す。

定義 2：連想記憶 (associative memory) 入力パターンの集合 I 、出力パターンの集合 O に対してニューラルネットワーク n が写像 $n: I \rightarrow O$ として定まることである。また $I = O$ であるような連想記憶を自己相関連想記憶 (autoassociative memory)、 $I \neq O$ であるような連想記憶を相互相関連想記憶 (heteroassociative memory) という。

代表的な相互結合型ニューラルネットワークであるホップフィールドモデルは最適化問題にも用いられている。

3.3 ホップフィールドモデル

3.3.1 ホップフィールドモデル (Hopfield model)

ホップフィールドモデルはアメリカの物理学者 J.J. Hopfield が導入したニューラルネットワークである。以下にホップフィールドモデルの特徴を示す。

- 複数のニューロンで構成され、各ニューロンは自分以外の全てのニューロンから出力を受け取る
- エネルギー関数の極小化するように各ニューロンの内部状態が変化する
- ニューロン間のシナプス結合荷重が対称である
- ニューロンの出力は各ニューロンの内部状態に依存する
- ある時刻において、1つのニューロンのみが他のニューロンからの出力を受け内部状態変化を起こす

ホップフィールドモデルは当初 2 値の出力値を扱うものとして提案されたが、その後連続値へと拡張された。このモデルは物理モデルから発展したものであり、脳のモデル化と言い難いものである。したがって、あくまでニューラルネットワークの特徴の解析から発生した 1 つの数理モデルであるにとらえるべきである。

ニューラルネットワークの状態は各ニューロンの出力を並べてベクトルとすることで表わすことができる。

3.3.2 2 値ホップフィールドモデル

本節では $n \in \mathbb{N}$ 個の 2 値出力のニューロンからなるホップフィールドモデルを考える。離散時間 t ($t = 0, 1, 2, \dots$) のときにニューロン i ($i = 1, 2, \dots, n$) の内部状態 $u_i(t)$ 、出力値 $x_i(t)$ を

$$x_i(t+1) = \begin{cases} 1, & u_i(t) > 0 \\ x_i(t), & u_i(t) = 0 \\ 0, & u_i(t) < 0 \end{cases} \quad (3.1)$$

$$u_i(t) = \sum_{j=1}^n w_{i,j} x_j(t) - \theta_i \quad (3.2)$$

と定義する。ここで $w_{i,j}$ はニューロン i とニューロン j 間のシナプス結合荷重、 θ_i はニューロン i の時刻 t におけるしきい値である。また $w_{i,j} = w_{j,i}$, $w_{i,i} = 0$ である。

式 (3.2) は神経細胞モデルにおいて膜電位の計算式と同等であるため、式 (3.1) とあわせてみると、膜電位を越えたら出力は 1 となり膜電位を越えない場合は出力は 0 となるととらえることができる。ホップフィールドモデルにおいてニューロンの内部状態とは階層型ニューラルネットワークでの膜電位と同じ意味の単語である。膜電位ではなく内部状態という単語を用いているのは、ホップフィールドモデルが物理から発展したモデルであるためである。

ニューラルネットワークの状態は各ニューロンの出力を並べたものであるため、 n 個のニューロンによる 2 値ホップフィールドモデルならば 2^n の状態を持つ。

このニューラルネットワークのエネルギー関数 E は

$$E = -\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n w_{i,j} x_i x_j + \sum_{i=1}^n \theta_i x_i \quad (3.3)$$

として定義される。式 (3.3) の値は任意の状態から式 (3.1) と式 (3.2) に従って内部状態変化を繰り返すことで必ず平衡状態へと遷移する。

Proof. ホップフィールドモデルにおいて 1 度の内部状態変化では 1 つのニューロンの内部状態のみが変化する。よって 1 度の内部状態変化でエネルギー関数の値が減少もしくは変らないことを示せばよい。

まず式 (3.3) をある k ($k = 1, 2, \dots, n$) を用いて

$$\begin{aligned} E &= -\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n w_{i,j} x_i x_j + \sum_{i=1}^n \theta_i x_i \\ &= -\frac{1}{2} \sum_{i \neq k} \sum_{j \neq k} w_{i,j} x_i x_j + \sum_{i \neq k} \theta_i x_i - \frac{1}{2} x_k \sum_{i=1}^n w_{i,k} x_i - \frac{1}{2} x_k \sum_{j=1}^n w_{k,j} x_j + \theta_k x_k \end{aligned}$$

と変形すると $w_{i,j} = w_{j,i}$ であるため

$$E = -\frac{1}{2} \sum_{i \neq k} \sum_{j \neq k} w_{i,j} x_i x_j + \sum_{i \neq k} \theta_i x_i - x_k \sum_{i=1}^n w_{i,k} x_i + \theta_k x_k \quad (3.4)$$

となる。

ここである t ($t = 0, 1, 2, \dots$) においてニューロン k が式 (3.1) と式 (3.2) に従って内部状態変化を起こしたとすると、 k でない任意の i ($i = 1, 2, \dots, n$) に対して

$$\begin{aligned} x_k(t+1) &= x_k(t) + \Delta x_k \\ x_i(t+1) &= x_i(t) \end{aligned}$$

が成立する。ここで $\Delta x_k \in \{-1, 0, 1\}$ である。式 (3.4) から時刻 t と $t+1$ のエネルギー $E(t), E(t+1)$ を計算するとそれぞれ

$$\begin{aligned} E(t) &= -\frac{1}{2} \sum_{i \neq k} \sum_{j \neq k} w_{i,j} x_i(t) x_j(t) + \sum_{i \neq k} \theta_i x_i(t) - x_k(t) \sum_{i=1}^n w_{i,k} x_i(t) + \theta_k x_k(t) \\ E(t+1) &= -\frac{1}{2} \sum_{i \neq k} \sum_{j \neq k} w_{i,j} x_i(t) x_j(t) + \sum_{i \neq k} \theta_i x_i(t) - x_k(t+1) \sum_{i=1}^n w_{i,k} x_i(t) + \theta_k x_k(t+1) \end{aligned}$$

となり、エネルギー関数の差分 ΔE は

$$\begin{aligned} \Delta E &= E(t+1) - E(t) \\ &= -[x_k(t+1) - x_k(t)] \sum_{i=1}^n w_{i,k} x_i(t) + [x_k(t+1) - x_k(t)] \theta_k x_k(t+1) \\ &= -\left[\sum_{i=1}^n w_{i,k} x_i(t) - \theta_k \right] \Delta x_k \end{aligned}$$

と書くことができ、式 (3.2) から

$$\Delta E = -u_k(t) \Delta x_k$$

が得られる。時刻 t から $t+1$ にかけてニューロン k の内部状態が変化しなかった場合は $\Delta x_k = 0$ であるため $\Delta E = 0$ となる。内部状態が変化した場合を考えると $u_k(t) > 0$ かつ $x_k(t) = 0$ というパターンと $u_k(t) < 0$ かつ $x_k(t) = 1$ というパターンが考えられる。 $u_k(t) > 0$ かつ $x_k(t) = 0$ である場合、式 (3.1) から $x_k(t+1) = 1$ であることため $\Delta x_k = 1$ となり $\Delta E < 0$ となる。 $u_k(t) < 0$ かつ $x_k(t) = 1$ である場合も、式 (3.1) から $x_k(t+1) = 0$ であることため $\Delta x_k = -1$ となり $\Delta E < 0$ となる。よってどのような場合であっても $\Delta E \leq 0$ が成立する。したがって任意の t ($t = 0, 1, 2, \dots$) に対してエネルギーの差分 $\Delta E = E(t+1) - E(t)$ は $\Delta E \leq 0$ となり、 $E(t)$ を時間の関数とみると単調減少であることが言える。

また、エネルギー関数の値はニューラルネットワークの状態に対して一意に定まるため同じニューラルネットワークの状態に遷移することはない。さらにニューラルネットワークの状態が有限であることから、必ずエネルギー関数の値に平衡点が存在し、エネルギー関数の値は時間の増加に対して収束する。□

3.3.3 連想記憶への応用

2 値ホップフィールドモデルは連想記憶に適用された。これを実現する方法として、不正確なパターンを入力とし、ニューラルネットワークの状態遷移を発生させ、最終的に到達したニューラルネットワークの状態を想起パターンとすることである。また以降 2 値ホップフィールドの出力値は 0,1 でなく $-1,1$ を用いる。

パターンをベクトルで表現し各成分を各ニューロンの出力値とすれば、パターンのベクトル表記はニューラルネットワークの状態とみることができる。この記憶させることを記録させるという。 P 個のパターンを記録させようとし、記録させたい s ($s = 1, 2, \dots, P$) 個目のパターンを $\mathbf{x}^s \in \{-1, 1\}^n$ と表わすと、ニューラルネットワークにパターン \mathbf{x}^s を記録させるということは、エネルギー関数が \mathbf{x}^s で極小値をとるようにシナプス結合荷重を与えることと等しくなる。ここで、各ニューロンのしきい値が 0 である場合のシナプス結合荷重の設計を考える。

各ニューロンのしきい値を 0 とするとエネルギー関数 (式 (3.3)) は

$$E = -\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n w_{i,j} x_i x_j$$

となる。ここで、 P 個のパターン $\mathbf{x}^s = (x_1^s, x_2^s, \dots, x_n^s) \in \{-1, 1\}^n$ を極小値とするシナプス結合荷重の一つは

$$w_{i,j} = \sum_{s=1}^P x_i^s x_j^s \quad i, j = 1, 2, \dots, n$$

である。これは、この式をエネルギー関数に代入することで

$$\begin{aligned}
E &= -\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \sum_{s=1}^P x_i^s x_j^s x_i x_j \\
&= -\frac{1}{2} \sum_{s=1}^P \sum_{i=1}^n \sum_{j=1}^n x_i^s x_i x_j^s x_j \\
&= -\frac{1}{2} \sum_{s=1}^P \sum_{i=1}^n x_i^s x_i \sum_{j=1}^n x_j^s x_j \\
&= -\frac{1}{2} \sum_{s=1}^P \left(\sum_{i=1}^n x_i^s x_i \right)^2 \\
&= \frac{1}{2} \sum_{s=1}^P \left[- \left(\sum_{i=1}^n x_i^s x_i \right)^2 \right]
\end{aligned}$$

となり、 x_i^s や x_i が -1 か 1 しかとらないことから、エネルギー関数が各 \mathbf{x}^s を極小値とする 2 次関数の和となっていることがわかる。パターンの数 P が少ない等いくつかの制約の上ではこのエネルギー関数が任意の \mathbf{x}^s を極小値とすることが直感的にわかる。また、 \mathbf{x}^s と $-\mathbf{x}^s$ はエネルギーの値が同じであるため \mathbf{x}^s を記録することと $-\mathbf{x}^s$ を記録することは同じ意味を持つ。

なお、大きな n 似たいして記録ベクトルをランダムに選ぶことで $\{-1, 1\}$ の 2 値ホップフィールドモデルの記憶容量は

$$\frac{n}{2 \log_2 n}$$

で近似できることが示されている。

3.3.4 連続値ホップフィールドモデル

本節では連続値ホップフィールドモデルについて考える。連続値ホップフィールドモデルと 2 値ホップフィールドモデルとの違いは、出力値がとりうる値の区間が $[0, 1]$ であることである。また、出力値の変化に共なってニューロン i の内部状態 $u_i(t)$ の変化は微分方程式

$$\frac{du_i(t)}{dt} = -\frac{u_i(t)}{\tau} + \sum_{j=1}^n w_{i,j} x_j(t) - \theta_i \quad (3.5)$$

$$x_i(t) = f_i(u_i(t)) \quad (3.6)$$

で表わされる。ここで $i = 1, 2, \dots, n$ であり、 $f_i(u)$ は非線形な連続単調増加関数である。また、 τ はニューロンの内部状態における時定数である。

式 (3.5) の第 1 項目は $u_i(t)$ を指数関数的に減少させることで、 $u_i(t)$ の発散を抑制し平衡値へと近づける効果がある。第 1 項を除けば式 (3.5) は式 (3.1) と同等の式とみることができ、膜電位に対応する値が正であれば $u_i(t)$ は時間的に増加させる、負であれば $u_i(t)$ は時間的に減少させる、0 であれば $u_i(t)$ の変化に寄与しないという効果がある。また

$$u_i(t) = \tau \left(\sum_{j=1}^n w_{i,j} x_j(t) - \theta_i \right)$$

のとき式 (3.5) は 0 となり平衡状態となり、 $\tau = 1$ のとすると 2 値ホップフィールドモデルの式 (3.2) と一致する。

このネットワークに対してエネルギー関数は

$$E = -\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n w_{i,j} x_i x_j + \sum_{i=1}^n \theta_i x_i + \frac{1}{\tau} \sum_{i=1}^n \int_0^{x_i} f_i^{-1}(x) dx \quad (3.7)$$

と定義される。ここで $f_i^{-1}(x)$ は $f_i(x)$ の逆関数である。式 (3.7) は 2 値ホップフィールドモデルのエネルギー関数 (式 (3.3)) に項を追加したものである。これはエネルギー関数が時間 t に対して単調減少関数であるために追加された項である。

2 値ホップフィールドモデルと同様に時間変化によるエネルギーの変化を考える。エネルギー関数 (式 (3.7)) を時間 t によって微分すると式 (3.4) によって

$$\begin{aligned} \frac{dE}{dt} &= \sum_{i=1}^n \frac{\partial E}{\partial x_i} \frac{dx_i}{dt} \\ &= \sum_{i=1}^n \left(-\sum_{j=1}^n w_{i,j} x_j + \theta_i - \frac{1}{\tau} f_i^{-1}(x_i) \right) \frac{dx_i}{dt} \\ &= -\sum_{i=1}^n \left(\sum_{j=1}^n w_{i,j} x_j - \theta_i + \frac{1}{\tau} f_i^{-1}(x_i) \right) \frac{dx_i}{dt} \end{aligned}$$

となり、 f_i^{-1} が f_i の逆関数であることと式 (3.5) と式 (3.6) から

$$\frac{dE}{dt} = \sum_{i=1}^n \frac{du_i}{dt} \frac{dx_i}{dt}$$

となり、また

$$\frac{dx_i}{dt} = \frac{df_i}{du_i} \frac{du_i}{dt}$$

であるため

$$\frac{dE}{dt} = \sum_{i=1}^n \frac{df_i}{du_i} \left(\frac{du_i}{dt} \right)^2$$

が得られる。よって $\frac{df_i}{du_i} > 0$ であることから

$$\frac{dE}{dt} \leq 0$$

となり、エネルギー関数 (式 (3.7)) が時間に対して単調減少であることが示された。エネルギー関数が単調減少であるため、連続値ホップフィールドモデルのエネルギーは時間 $t \rightarrow \infty$ によって極小値となる。

ホップフィールドは単調増加関数としてシグモイド関数

$$f_i(u_i) = \frac{1}{1 + e^{-\frac{2u_i}{\mu_0}}}$$

を用いており、ここで μ_0 は基準活性化レベル (reference activation level) と呼ばれる。

3.3.5 最適化問題への応用

組み合わせ最適化問題 (combinatorial optimization problem) は、目的関数を最小化する組み合わせを探す問題である。ここで組み合わせ最適化問題の変数列をベクトル $\mathbf{x} \in \{0, 1\}^n$ へ変換し、また目的関数と式 (3.3) が一致するような $w_{i,j}$ を求めることができれば、2 値ホップフィールドモデルの平衡状態は組み合わせ最適化問題の目的関数の極小値と同等である。よって 2 値ホップフィールドモデルは組み合わせ最適化問題を解きうるモデルであるといえる。本節では NP 困難である組み合わせ最適化問題の中で代表的な問題である TSP を解くことを考える。

3.3.6 連続値ホップフィールドモデルの改良

3.4 ボルツマンマシン

3.4.1 ボルツマンマシンの動作

3.4.2 ボルツマンマシンの学習

3.4.3 ボルツマンマシンの特徴

4 競合学習型ニューラルネットワーク

4.1 認識機構の自己形成

4.2 生体のトポロジカルマッピングのモデル

4.3 コホーネンのモデル

4.3.1 予備実験

4.3.2 特徴抽出細胞の形成

4.3.3 コホーネンの学習則

4.3.4 コホーネンの自己組織化特徴マップのアルゴリズムとシミュレーション

4.3.5 応用例

5 ニューラルネットワーク研究の意義

5.1 特徴を生かす

5.1.1 研究の歴史

5.1.2 生物内のニューラルネットワークと人工ニューラルネットワーク

5.1.3 シナプスの可塑性と脳・神経系の可塑性

5.1.4 教師あり学習と教師なし学習

5.1.5 ニューロンコンピュータ

5.1.6 融合化技術

5.2 応用

5.2.1 応用されてきた分野

5.2.2 事例の完備性と適用有効範囲

5.2.3 ブラックボックスモデルの利用環境への適合性

5.3 脳科学への貢献