

誤差逆伝搬学習における局所解の吸引域の形状について (XOR 問題の場合)

正員 堀川 洋†

Shapes of Basins of Local Minima in a Backpropagation Network for the XOR Problem

Yo HORIKAWA†, Member

† 香川大学教育学部情報科学科, 高松市
Faculty of Education, Kagawa University, Takamatsu-shi, 760 Japan

あらまし 誤差逆伝搬型ニューラルネットワークの荷重空間における局所解の吸引域の形状を, 計算機シミュレーションにより調べた. そして, 簡単な XOR 問題においても, 局所解の吸引域は複雑な形状をしていることを示した.

キーワード ニューラルネットワーク, バックプロパゲーション, ローカルミニマム, 初期値設定問題

1. まえがき

誤差逆伝搬法を用いた層状ニューラルネットワークの学習では, 荷重の初期値によっては局所解(ローカルミニマム)に陥る. ここでは, 荷重空間における局所解の吸引域, すなわち, 局所解に陥る初期値の領域を, 排他的論理和 (XOR) の学習問題の場合について, 計算機シミュレーションにより調べた結果を示す.

2. モデル

図 1 に示す 2-2-1 型の 3 層ネットワークを用いて, 誤差逆伝搬法による XOR 問題の学習を行う. 入力信号: $x^a=(x_1^a, x_2^a)$ と教師信号 d^a を,

$$x^1=(0,0), d^1=0; x^2=(1,0), d^2=1;$$

$$x^3=(0,1), d^3=1; x^4=(1,1), d^4=0;$$

とし, 9 個の荷重: w を次式に従って学習させる (学習の収束条件は, $E(x(t)) < 0.01$ とする).

$$w(t+1) = w(t) - \partial E(w(t)) / \partial w(t)$$

$$w = (w_{01}, w_{02}, w_{11}, w_{12}, w_{21}, w_{22}, v_0, v_1, v_2)$$

$$E(w(t)) = 1/8 \sum_{a=1}^4 (o(t)^a - d^a)^2 \quad (1)$$

このネットワークの局所解における出力: o^a は, 次の四つの型に限られることが知られている⁽¹⁾.

(A) $o^1=o^2=o^3=o^4=1/2$

(B) $o^1=o^2=1/2, o^3=1, o^4=0$

$$(o^1 \neq o^4, o^2 \neq o^3)$$

(C) $o^1=o^2=o^3=2/3, o^4=0 (o^1 \neq o^4)$

(D) $o^1=o^3=o^4=1/3, o^2=1 (o^2 \neq o^3)$

これらのうち, (A) のものは, 原点: $w=0$ を含むため,

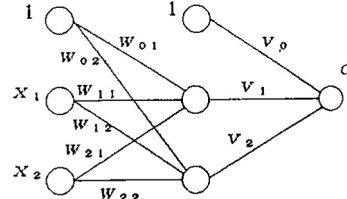
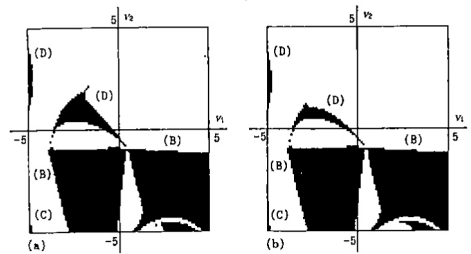


図 1 ネットワークの構成

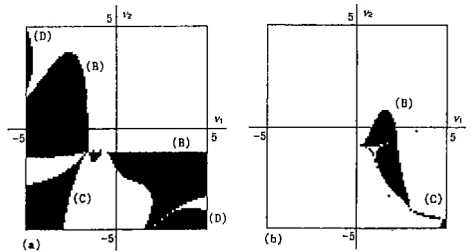
Fig. 1 Architecture of a 2-2-1 network for the XOR problem.



(a) $N=10^4$, (b) $N=10^5$

図 2 局所解の吸引域 ($|w(0)| \sim O(1)$)

Fig. 2 Basins of local minima in the $v_1(0)-v_2(0)$ plane. $w(0)=(0.3, -0.5, -1.4, 0.8, 0.6, -2.0, 1.0, *, *)$.



(a) $w(0)=(0.3, -0.5, -1.4, 0.8, 0.6, -2.0, -1.0, *, *)$

(b) $w(0)=(0.3, -0.5, 0.6, 0.8, 0.6, -2.0, 1.0, *, *)$

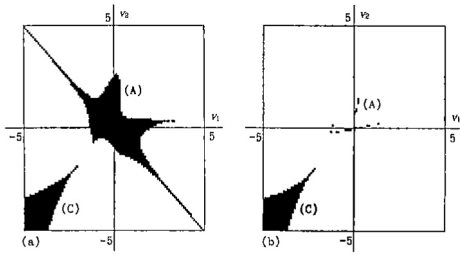
図 3 局所解の吸引域

Fig. 3 Basins of local minima ($N=10^5$).

局所解ではなくサドル型のものであると考えられる. それに対して, (B), (C), (D) は, $|w| = \infty$ となる局所解に対応している.

3. 局所解の吸引域

図 2, 3, 4 に, 計算機シミュレーションにより得た局所解の吸引域を示す. いずれも, 9 個の荷重のうち, 中間層-出力層間の 2 個の荷重の初期値: $v_1(0), v_2(0)$ について $[-5, 5]$ の範囲で (他の 7 個の荷重の初期値は同じにして), 学習の収束を調べたものである. 図中で黒色の領域が, N 回反復後においても学習が収束しなかつ



(a) $N=10^4$, (b) $N=10^5$

図4 局所解の吸引域 ($|\mathbf{w}(0)| \sim O(10^{-1})$)

Fig. 4 Basins of local minima.

$\mathbf{w}(0) = (-0.02, -0.04, -0.03, 0.04, 0.05, -0.09, -0.08, *, *)$.

た初期値であり、局所解の吸引域を含んでいる。但し、そのうちには、大域解の吸引域ではあるが N 回では学習が収束しなかった(収束の遅い)領域も含まれる(当然ながら、有限回のシミュレーションでは、局所解の吸引域と収束の遅い領域とを区別することはできない)。

図2は、 $|\mathbf{w}(0)| \sim O(1)$ の場合であり、(B)、(C)、(D)の三つの型の局所解の吸引域が存在している。ここで、図の中央付近にある弧状の領域において、反復回数を増やすと ($N: 10^4 \rightarrow 10^5$) 消滅した部分は、収束の遅い領域であるが、その途中では(D)の出力値をとっている。

また、図3は、図1のものと荷重の初期値を1個だけ変えた場合の結果である。吸引域の形状は、大きく

異なっている。

一方、図4は、初期値として、より小さな値 ($|\mathbf{w}(0)| \sim O(10^{-1})$) を用いたものであり、(A)の出力値をとる領域が見られる。但し、(A)の解はサドル型であるため、 $N=10^5$ ではほとんどすべてが大域解に収束する。

4. むすび

誤差逆伝搬学習における局所解の吸引域は、XOR問題を解くような小さなネットワークにおいても、ロールシャッハテストの図形を連想させる複雑で多様な形状をもつことを示した。このような吸引域の複雑さは、初期値設定問題の困難さに対応している。

誤差関数 ($E(\mathbf{w})$) の荷重空間における形状は、その1次元あるいは2次元断面を見る限りでは、平坦な丘と谷からなる滑らかなものであることが知られている⁽²⁾。そのため、吸引域の複雑さは、系の高次元性(この場合9次元にすぎないが)によるものであり、ネットワークの規模の増大に伴いその複雑さも増し、初期値設定がより困難なものとなることを示唆する。

文 献

- (1) Lisboa P. J. G. and Perantonis S. J.: "Complete solution of the local minima in the XOR problem", *Network*, 2, pp. 119-124 (1991).
- (2) 郷原一寿, 内川嘉樹: "階層型ニューラルネットワークにおける学習曲面の解析", 信学技報, NC90-43 (1990).
(平成5年5月26日受付, 6月30日再受付)