

Gauss-Sigmoid ニューラルネットワーク

大分大学 工学部 柴田克成、前原伸一、杉坂政典
東京工業大学大学院 知能システム科学専攻 伊藤宏司

Gauss-Sigmoid Neural Network

Katsunari Shibata (shibata@cc.oita-u.ac.jp), Shinnichi Maehara and Masanori Sugisaka, Oita University
Koji Ito, Tokyo Institute of Technology

Abstract- Recently RBF(Radial Basis Function)-based networks have been widely used because they can learn a strong non-linear function faster and more easily by their local learning characteristics. However, it has no hidden units that can represent some global information. Accordingly even if the knowledge obtained through the previous sets of learning is utilized effectively in the present learning, the network has to learn from scratch. Then the present paper proposes a Gauss-Sigmoid neural network for learning with continuous input signals. The input signals are put into a RBF network, and then the outputs of RBF network are put into a sigmoid-based multi-layered neural network. After learning based on back-propagation, the localized signals by the RBF network are integrated and an appropriate space for the given learning is reconstructed in the hidden layer of the sigmoid-based neural network. The advantage was examined by the simulation of a supervised learning task that requires a discontinuous mapping. It is also confirmed that the Gauss-Sigmoid neural network shows better performance than sigmoid-based neural networks in the reinforcement learning task in which instability was shown when the sigmoid-based neural network was utilized.

Keywords- RBF ネットワーク, Gaussian soft-max ネットワーク, sigmoid-based ニューラルネット,
Gauss-sigmoid ニューラルネット, 局所化, 内部表現.

1. はじめに

近年、非線形関数近似を行う際に、シグモイド型ニューラルネット(NN)に代わり、Fig. 1 (a)のような RBF (Radial Basis Function) ネットワークや、RBF ネットの汎化能力を改善した Gaussian soft-max ネットワークが使われることが多い。情報の局所化を行うガウシアン(RBF)ベースのネットワークは、シグモイド型 NN と比較し、学習が速い、複雑な関数を容易に近似できるということがその理由としてあげられる。

強化学習の分野においても、Boyan らは、Hill-car 問題、2-Dpuddle world 問題などを例に、シグモイド型の NN を TD 型の強化学習に用いると学習が不安定になると報告している[1]。それに対し Sutton は、CMAC を用いると学習がうまくできることを示している[2]。また、Gordon は、k-nearest-neighbor のように、いくつかの代表点での値や定数の間の重み付き平均という形の関数近似の場合には強化学習が安定であり、線形近似やシグモイド NN では不安定になる場合があることを示している[3]。しかし、これらの情報を局所化するアプローチは、テーブルルックアップに近く、さまざまな欠点も持ち合わせている。

本稿では、RBF ベースのネットワークとシグモイド型 NN を比較し、利点と欠点を整理するとともに、両者の利点を引き出すために、Fig. 1(b)のように、ガウシアンを通して局所化した信号をシグモイド型 NN に入力して処理する Gauss-Sigmoid ニューラルネットワークを提案する[4]。そして、その有効性を教師あり学習のシミュレーションを通して示すとともに、Boyan らが指摘した強化学習の例題を Gauss-Sigmoid NN を使って解くことによって、連続値入力を直接シグモイド型 NN に入力する場合と比較して、より良い解が得られることを示す[5]。

2. RBF(Gauss)-based ネット対シグモイド型 NN

2.1 RBF ネットと Gaussian soft-max ネット

最初に、RBF ネットと汎化能力を改善した Gaussian soft-max ネットを紹介する。RBF ネットでは、ガウシア

ン $g(x)$ を用いており、ネットワークの出力 o は、

$$o = \sum_{i=1}^n w_i g_i(\mathbf{x}) + \theta, \quad \text{where} \quad (1)$$

$$g_i(\mathbf{x}) = \exp\left(-\frac{1}{2} \sum_{d=1}^D \frac{(x_d - \mu_{i,d})^2}{\sigma_{i,d}^2}\right) \quad (2)$$

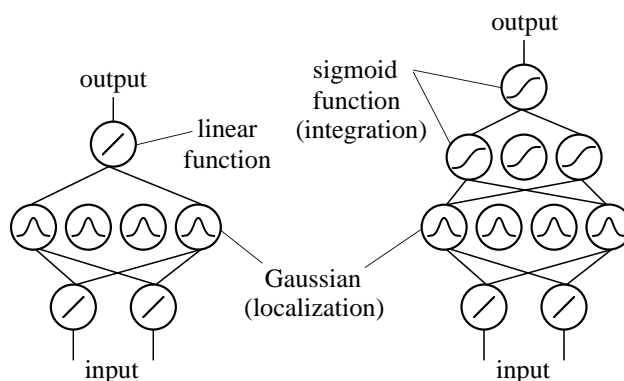
と書くことができる。ただし、 \mathbf{x} : 入力パターン、 $(\mu_{i,1}, \dots, \mu_{i,D})$: i 番目の RBF ユニットの中心、 $(\sigma_{i,1}, \dots, \sigma_{i,D})$: i 番目のガウシアンサイズ、 w_i : i 番目のガウシアンから出力への結合の重み値、 θ : バイアス、 n : ガウシアンの数、 D : 入力の次元である。

一方、Gaussian soft-max ネットでは、各ガウシアンユニットの出力は、すべてのガウシアンの出力の和で正規化されるため、その出力 o は、

$$o = \sum_{i=1}^n w_i b_i(x) + \theta, \quad \text{where} \quad (3)$$

$$b_i(x) = g_i(x) / \sum_{j=1}^n g_j(x) \quad (4)$$

と表される。



(a) RBF network (b) Gauss-Sigmoid NN
Fig. 1 RBF network and Gauss-Sigmoid neural network.

ガウシアンユニットの中心 μ およびサイズ σ は、誤差逆伝搬(BP)法で学習することができる。ただし、ガウシアンユニットのサイズ σ が小さくなると、急峻な形状となるため、 μ および σ の更新量が大きくなり過ぎてしまう。そこで、ここでは、 σ を \log スケールで扱うとともに、最急勾配法によって求められたそれぞれの更新量に σ をかけることによって更新量を調節した。具体的には、伝搬誤差を δ として、

$$\Delta\mu_{j,d} = g_j(\mathbf{x}) \frac{x_d - \mu_{j,d}}{\sigma_{j,d}} \delta_j \quad (5)$$

と中心 μ を更新するとともに、サイズ σ は、

$$\sigma_{j,d} = \exp(s_{j,d}) \quad (6)$$

の式で規定される s を

$$\Delta s_{j,d} = g_j(\mathbf{x}) \frac{(x_d - \mu_{j,d})^2}{\sigma_{j,d}^2} \delta_j \quad (7)$$

と更新して、再び(6)式により s を σ に変換して用いた。

2.2 学習速度

RBF 型ネットでは、各 RBF、つまり、ガウシアンユニットの出力は、入力パターンがその中心から離れると 0 に近づく。したがって、ある入力パターンでの学習が、他の離れた入力パターンに対する出力に影響を及ぼさないため、非線形性の強い関数関係を学習する場合には一般的に学習は速い。一方、シグモイド型 NN は、各ニューロンの出力関数であるシグモイド関数が、線形に近い関数であるため、ある入力パターンでの誤差による学習が、離れた入力パターンにも影響を及ぼす。そのため、線形に近い関数近似の場合は、速く学習ができるが、非線形性が強い場合は、すでに学習した関数関係を壊してしまい、学習が遅くなったり不安定になる場合が多い。

しかし、RBF ネットにおいて、多数のガウシアンを割り当てた場合、学習はそれだけ遅くなる。一方、シグモイド型 NN の中間層ニューロン数を増やしても、学習速度はあまり変化はないか、むしろ、速くなる場合が多い。以上より、非線形性が強い関数関係を近似する場合は RBF ネットが、そうでない場合はシグモイド型 NN が学習速度においては有利であると言える。

2.3 汎化能力

通常の RBF ネットは、前述のように、各ガウシアンが、入力空間の一部の情報しか表現していないため、テーブルルックアップよりは汎化能力が優れているものの、シグモイド型 NN と比較して大きく劣る。しかし、RBF ネットワークの拡張である Gaussian soft-max ネットでは、各ガウシアンユニットの相対的な出力値によって出力が決定されるため、近傍にガウシアンが配置されていないような場合に、外挿がきくという点で通常の RBF ネットよりも汎化能力に優れていると言える。たとえば、Fig. 2 のように、2次元の入力空間を考え、各象限ごとに教師信号が異なる場合を考える。この場合、原点に近い部分に4つのガウシアンを配置し、原点近傍だけで学習を行うと、点 α のように、原点から離れてガウシアンが配置されていない部分では、RBF ネットの場合、出力はほぼ 0 になるが、Gaussian soft-max ネットの場合には、教師信号に近い値が出力されるようになる。

しかしながら、ガウシアンが密に配置されているところでは、Gaussian soft-max ネットも汎化能力が劣る。たとえば、Fig. 3 のように、二次元の入力空間を2つの領域に

線形分離をする場合に、分散の小さい6つのガウシアンユニットが配置されていたとすると、RBF ネットの場合も Gaussian soft-max ネットの場合も、両端のガウシアンユニットが学習しても、真中のユニットが学習しない限り、十分に補間されない。一方、シグモイド型 NN はより強い汎化能力を有しており、Fig. 2 や Fig. 3 のような場合でも補間されやすい。

2.4 内部表現

シグモイド型の多層ニューラルネットでは、中間層に内部表現が形成される。しかし、RBF ネットでは、局所的な情報を表すガウシアンのみで、大域的な情報を表現できる中間層がない。Fig. 3 の例で、入出力関係がすでに学習され、内部表現をシグモイド型 NN が獲得していたとすると、もう一つの出力ニューロンを設けて、中間層と出力層の間の結合の重み値を 0 としても、領域 A で学習すれば領域 B での出力も正しい出力に近くなると予想される。しかし、RBF ネットの場合には、すべての領域で正しい出力を得るためには、すべてのガウシアンユニットが配置されている領域で学習をしなければならない。

このような能力は、単一の関数関係を学習させる場合には特に必要とはならない。しかし、ロボットの学習において、たとえば視覚センサ信号から三次元の空間認識を行うことは、あらゆるタスクで必要となる。しかし、このような場合でも RBF ネットで学習すると、タスクが変わると一から学習し直さなければならない。一方、シグモイド型 NN では、中間層に三次元空間認識の情報が保持され、次のタスクの学習が加速されると考えられる。

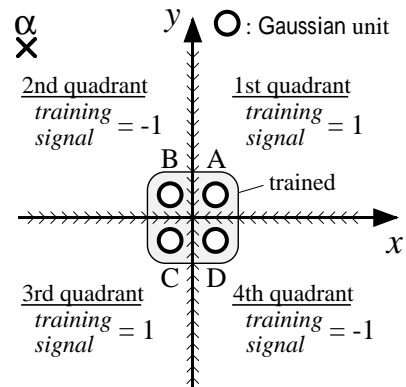


Fig. 2 An example which indicates the extrapolation ability of Gaussian soft-max network in two dimensions of input space using 4 RBF units.

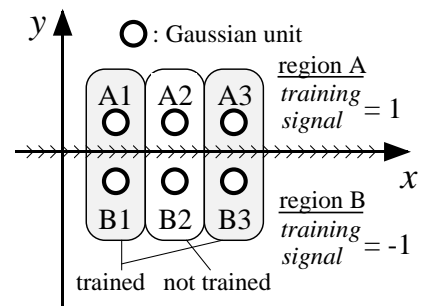


Fig. 3 An example case in which the generalization is not effective in Gaussian soft-max network.

また、逆に、三次元空間認識を必要とするさまざまなタスクを経験することにより、二次元の視覚情報から三次元空間を中間層の中で再構成することができるように考えられる。そして、このような能力こそが知能の形成に大きな役割を果たしていると考えられる。

2.5 不連続関数の近似

連続値入力に対し出力が不連続な不連続関数は、典型的な強い非線形関数である。不連続関数を誤差 0 で近似するためには、ニューラルネット内の重み値が少なくとも一つは無限大でなければならない。しかし、実際にはそのようなことは不可能であるため、不連続部分では必ず大きな誤差が残ることになる。前述のように、シグモイド型 NN では、ある場所での誤差に基づく学習により、遠く離れた入力パターンにも影響がおよび、すでに学習したパターンを破壊し、学習が不安定になりやすい。また、不安定になることを避けて学習係数を小さくすると、学習が遅くなる。一方、RBF ネットでは、ガウシアンが局所的であるため、不連続部分での誤差の影響は、遠い入力パターンの場合にはおよばず、不連続部分の周辺に限られる。したがって、学習係数がある程度大きくても、入力空間全体が不連続部分の誤差のために不安定になることはない。さらに、局所的な表現をすることは、ちょうど虫めがねで小さい部分を見るように、情報を拡大する働きがあり、これも学習を容易にすると考えられる。

3. Gauss-Sigmoid アーキテクチャ

以上のように、シグモイド型 NN、RBF ネットともにそれぞれ利点と欠点を有する。RBF ネットは、連続入力空間を局所化することで学習の高速化、安定化につながる。一方、シグモイド型 NN は、大域的な情報を表現するとともに、中間層でその情報を保持することができる。つまり、RBF は入力側、シグモイドは出力側で有効である。そこで、両者の利点を引き出し、欠点をなくすため、Fig. 1(b)のように、入力信号を最初に RBF で処理し、その出力をシグモイド型 NN で処理することを提案する。これを、Gauss-Sigmoid ニューラルネット(NN)と呼ぶ。

また、生物において、視覚センサをはじめとする多くのセンサの受容野が局所的であり、その後パルス密度に基づいて活動を決定する神経細胞が情報を処理しているとすれば、まさに、生物においても、Gauss-Sigmoid NN が使われていると考えることもできる。

4. 教師あり学習のシミュレーション

本章では、RBF ネット、Gaussian soft-max ネット、シグモイド型 NN、Gauss-Sigmoid NN のそれぞれに、同じ教師あり学習を適用し、その結果を比較した。入力信号は 0 から 1 までの 2 つの連続値で、教師信号は、Fig.5 の一番左の図のように、円と斜めの帯状の二つの白い部分が 0.4、それ以外の黒い部分が -0.4 となっており、境界部分は不連続となっている。シグモイド関数の値域は -0.5 から 0.5 である。帯状の形は、シグモイドニューラルネットが比較的学習が得意であり、円は、RBF ネットが学習を得意とする。Gauss-Sigmoid NN のガウシアン数は 9 または 36、RBF および Gaussian soft-max ネットのガウシアン数は 36 とした。シグモイド NN は 5 層とし、入力層から 2-9-10-4-1 とした。Gauss-Sigmoid NN のシグモイド部分は、10-4-1 とし、Gauss-Sigmoid NN でガウシ

アンが 9 個の場合に、パラメータの数が、他の場合とほぼ同じになるようにした。各ネットワークの学習係数は、試行錯誤をもとに、誤差が小さくなるように、Table 1 のように設定した。また、シグモイド、ガウシアン学習には、慣性項を付加し、その係数を 0.9 とした。中間層の最上層 - 出力層間の結合の初期重み値をすべて 0 とし、その他の結合はすべて微小な乱数とした。また、各ガウシアン初期配置は、中心 μ を入力領域内に均等配分し、サイズ σ を隣接するガウシアン中心間距離と等しくした。ガウシアン数が 36 個の場合の初期配置は Fig. 6(a) のようになる。入力パターンは、毎回ランダムに決定した。50000 回学習後に学習係数を 10 分の 1 に減少させ、さらに、80000 回学習後に、出力層への重み値を 0 にリセットし、教師信号を -0.4 を 0.4 に、0.4 を -0.4 に入れ替えた。

Fig. 4 に、学習の進行による誤差の推移を 10 回のシミュレーションの平均値で示す。縦軸は、26x26 の格子上の 776 個の入力パターンに対する誤差の自乗和である。学習のもっとも初期には、RBF ネットの誤差が最も速く減少し、36 個のガウシアンを持つ Gauss-Sigmoid NN がそれに続いた。シグモイド NN が最も学習が遅い。しかし、その後、RBF ネットは、誤差が減らず、学習係数を減らした後は、逆に、誤差が最も大きくなる。一方、ガウシアンが 9 個の Gauss-Sigmoid NN は、最初はあまり誤差が減少しないが、その後、徐々に減少し、最終的には、ガウシアンが 36 個の Gauss-Sigmoid について誤差が小さ

Table 1 Learning rate

Network		(a) simulation 1	(b) simulation 2
Gauss-Sigmoid network	sigmoid	0.3	0.3
	gaussian	0.001	0.001
RBF network	linear	0.05	0.05
	gaussian	0.02	0.01
Gaussian Soft Max network	linear	0.05	0.05
	gaussian	0.01	0.0003
Sigmoid-based NN		0.3	0.3

- ① Gauss-Sigmoid (2-36-10-4-1)
- ② Gauss-Sigmoid (2-9-10-4-1)
- ③ RBF (2-36-1)
- ④ Gaussian Soft Max (2-36-1)
- ⑤ Sigmoid (2-9-10-4-1)

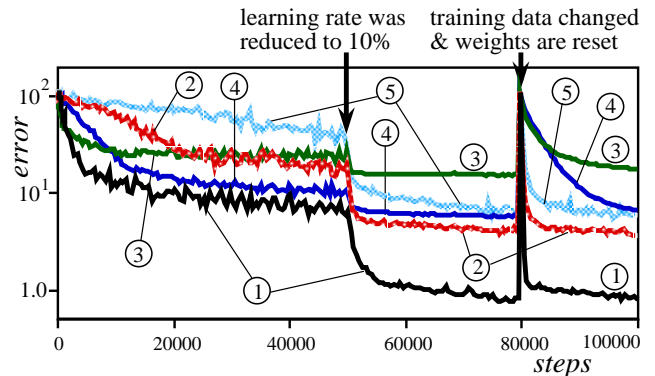


Fig. 4 Comparison of the learning curves between some networks. Gauss-Sigmoid NN with 36 RBF units learns fast and realize the best approximation.

くなった。Gaussian soft-max ネットは、学習の速度も、最終的な誤差の大きさも、中間的な性質を示している。

次に、80000 回の学習後に、教師信号を反転させ、出力層への結合を 0 にリセットした後の学習を見ると、シグモイドと Gauss-Sigmoid NN は、中間層に形成された内部表現を利用するため学習が速く、学習の初期に速かった RBF ネットは、学習係数が小さくなった分、逆に学習が遅くなっている。また、不連続な関数近似であるにもかかわらず、シグモイド NN の学習が安定であるのは、学習係数が十分小さいためで、その分、他と比較して学習が遅いと考えられる。

次に、中間層による内部表現の有効性を示すシミュレーションを行った。学習する問題は同じ問題とし、20000 回ごとに中間層の最上層と出力層間の結合の重み値をすべて 0 とするとともに、教師信号を反転させた。そして、この重み値のリセットと教師信号の反転を 48 回繰り返し、内部表現を形成させた。そして、再び重み値をリセットした後、原点と円の中心の 2 つの入力パターンのみ学習させ、未学習の入力パターンに対する出力を観察した。学習係数は、試行錯誤の結果、Table 1 (b) のようにした。

その結果を Fig. 5 に示す。一番左の図は教師信号を、上の 4 つの図は 48 回学習後の各ネットワークの出力分布を示している。また、下の 4 つは、2 つの入力のみ学習した後の出力分布を示している。学習した 2 つの入力パターンは、図中の小さな で示している。これより、学習後は、いずれも教師パターンに近い出力分布となっている。特に、Gauss-Sigmoid とシグモイド NN は、教師パターンとの誤差が小さいことがわかる。下の図を見ると、2 つの入力パターンを学習するだけで、Gauss-Sigmoid NN は、教師信号に非常に近い出力分布を実現している。また、シグモイド NN も、教師信号に比較的近い分布をしている。一方、RBF ネットは、円形部分の中心近くにあるガウシアンから出力ユニットへの結合係数が大きくなり、それ以外のガウシアンユニットからの結合の重み値はあまり大きくなりえないため、円の中心付近に中心を持つガウシアンの出力がそのままネットワークの出力となっていることがわかる。一方、Gaussian soft-max ネットの場合も、円の中心付近と、原点付近のガウシアンに対する重み値のみが学習されているため、やはり、帯状の部

分は期待通りの出力がされないことがわかる。ただし、Gauss-Sigmoid NN やシグモイド NN の場合でも、帯状の部分とその外側の部分の出力がほとんど同じとなってしまふなど、初期重み値の差や、学習時に乱数で選ぶデータの差によって、誤差が大きくなる場合もあった。しかし、いずれの場合も、RBF や Gaussian soft-max ネットと比較すると誤差は小さかった。また、シグモイドの層の数を増やし、上の層のニューロン数を少なくするほど誤差は小さくなる傾向にあった。一方、RBF および Gaussian soft-max ネットの場合は、ガウシアン初期分布は一定であるため、試行による誤差の差はほとんどなかった。

Fig. 6 に学習後のガウシアン中心とサイズの値を示す。図中の十字の交点がガウシアンの中心 μ を、そこからのため、横のそれぞれの長さが、それぞれの方向のサイズ σ を

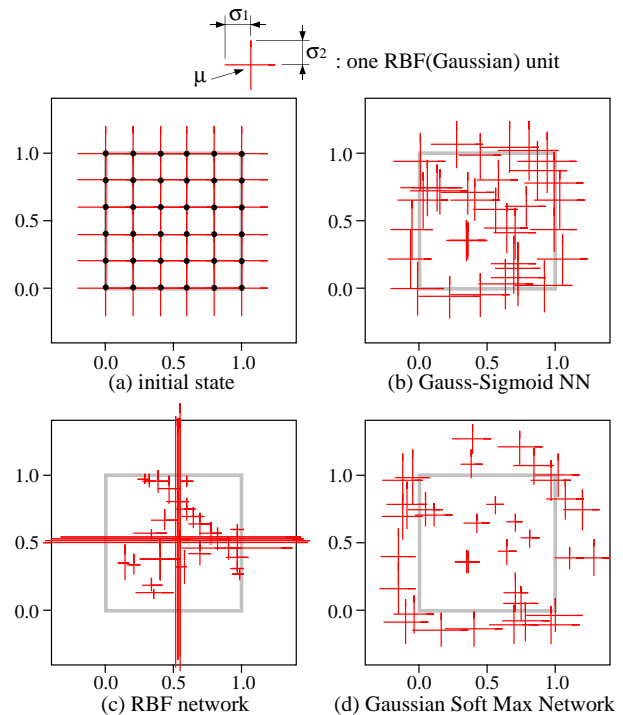


Fig. 6 The change of the center and size of each of 36 RBF units by learning for each network.

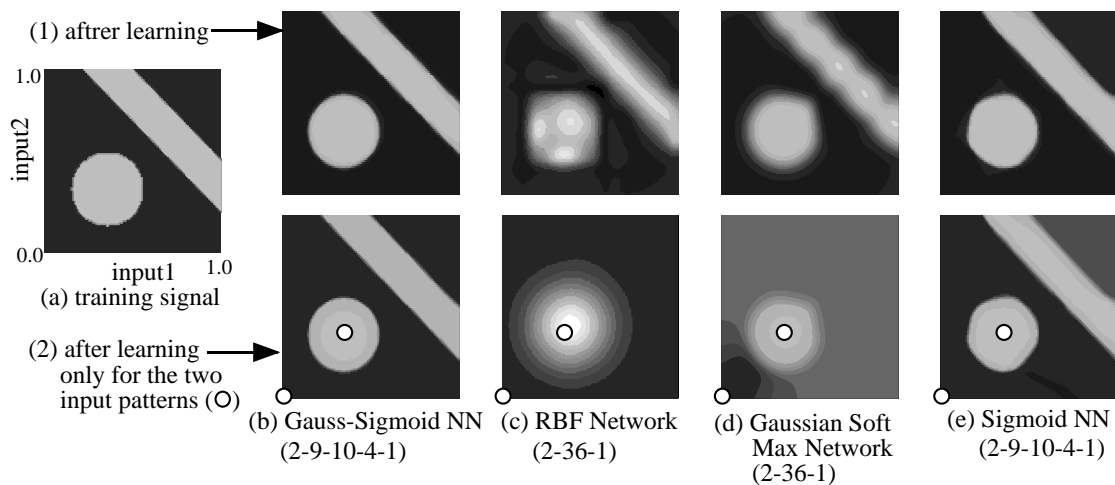


Fig. 5 Comparison of the output distribution after learning. The RBF-based networks cannot approximate correctly only by the learning of two input patterns because it does not have internal representation.

示す。ここでは、Gauss-Sigmoid NN は、ガウシアン 36 個の場合を示す。(a)は初期配置で、3つのネットワークいずれも同じである。この図より、RBF ネットは、入力空間の中の、円の中心や帯状の部分、およびその外側に存在していることがわかる。また、サイズが非常に大きいガウシアンがいくつか存在するが、これらはいずれも、出力層との結合の重み値は小さかった。一方、Gaussian soft-max ネットの場合は、入力範囲を大きくはみ出してガウシアンが存在していることがわかる。それに対し、Gauss-Sigmoid では、両者の中間的な分布で、ガウシアンが存在していることがわかる。また、Gauss-Sigmoid では、ガウシアンユニットの平均とサイズを変化させなくても、変化させた場合と比較して、誤差の大きさに大きな変化はなかった。これは、Gauss-Sigmoid では、ガウシানের後にシグモイド型のニューラルネットワークがあるため、局所化した入力を非常に柔軟に統合し、再構成することができるためと考えられる。

5. 強化学習のシミュレーション

前述のように、Boyan らは、シグモイド型のニューラルネットワークと強化学習の組み合わせが不安定を招く場合があることを指摘している。彼らは、実際に不安定になった問題の一つとして hill(mountain) car 問題を挙げている。本章では、この問題を Gauss-Sigmoid NN を用いて解き、他の場合と比較する。

5.1 Hill-car 問題

Hill(Mountain)-car 問題は、Fig. 7 のように、

$$\begin{cases} y = x(1+x) & \text{if } x < 0 \\ y = x/\sqrt{1+5x^2} & \text{if } x \geq 0 \end{cases} \quad (8)$$

と表される斜面を車が登り、 $x=1.0$ に到達すると報酬が、 $x=-1.0$ に来ると罰が与えられる。車の運動方程式は、

$$\frac{dv}{dt} = \left(\frac{\text{action}}{m} - g \frac{y'}{\sqrt{1+y'^2}} \right) / \sqrt{1+y'^2} \quad (9)$$

$$\frac{dx}{dt} = v, \quad y' = \frac{dy}{dx}$$

m : 車の質量、 action : 車の推力、 v : 速度

と記述できる。そして、斜面が急な部分では、推力を登る方向に最大にしても、重力によって斜面を下る方向に働く力の方が大きく、 $x=-0.5$ からスタートした場合、最大推力で右側の斜面を一気に登ろうとしても、初速度 0 で斜面を登りきることはできない設定となっている。これは、倒立振り子と似ている。したがって、初速度 0 で斜面を右に一気に登りきれぬかどうかの境界部分では、そこより右にいるときには右に、左にいるときには、左に推力を出して左側の斜面を登ってから再び戻ってきて右

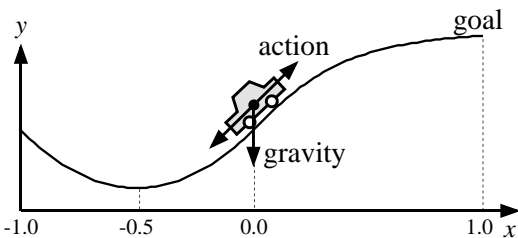


Fig. 7 Hill(Mountain)-car task.

側の斜面を登らなければならない。したがって、状態の評価値も、出すべき推力の大きさも不連続になる。また、左側に振りあがる場合は、あまり力を出しすぎると、左側の斜面からはみ出し、罰が与えられ、力が弱すぎると、次に右側の斜面に差し掛かったときに、登りきることができなくなる。ここでは、推力を -3.0 から 3.0 と制限したため、速度 0 で、 x が約 -0.74 より小さい状態を通らなければ左側の斜面を登ることができない。

5.2 シミュレーション

Hill-car 問題を actor-critic 型の強化学習によって学習した。その際、actor と critic を一つのネットワークで構成し、それを、シグモイド NN、Gauss-Sigmoid NN、RBF ネットの 3 つの場合で比較した。車の初期状態は、毎回 $-1 \leq x \leq 1, -4 \leq v \leq 4$ の範囲で乱数で決定した。そして、ネットワークの出力の一つである action に、微小な一様乱数を 3 乗した値 rnd を加えて実際の推力とする。そして、(9)式をルンゲクッタ法で解き、遷移後の状態を求め、車が丘の頂上に到達した場合は報酬 r を 1、その他の場合は 0 とした。ネットワークの出力は、状態評価値 P と車の推力 action の 2 つであり、TD 誤差

$$\hat{r}_t = r_t + \gamma P(x_t) - P(x_{t-1}) \quad (10)$$

を用いて、状態評価値 P は、

$$P_s(x_{t-1}) = P(x_{t-1}) + \hat{r}_t \quad (11)$$

を教師信号として学習し、推力 action は、

$$\text{action}_s(x_{t-1}) = \text{action}(x_{t-1}) + \alpha \hat{r}_t \text{rnd}_{t-1} \quad (12)$$

を教師信号として学習した。ただし、左側に飛び出した場合、および丘の頂上に達した場合は、 $P(x_t) = 0$ として学習した。速度 v は、 -4 より小さくなった場合は -4 、 4 より大きくなった場合は 4 とした。Gauss-Sigmoid およびシグモイド NN の中間層は 1 層とし、ニューロン数は 40 個とした。また、RBF ネットおよび Gauss-Sigmoid NN のガウシアンユニットの数は 110 個とし、 x 方向 10 個、 v 方向 11 個の格子の上に配置し、サイズが隣接ユニットの中心間距離になるようにした。ここでは、ガウシানের中心とサイズは学習させなかった。学習係数は、Gauss-Sigmoid NN が $100/\sqrt{\text{結合しているユニット数}}$ 、シグモイド NN が $80/\sqrt{\text{結合しているユニット数}}$ 、RBF ネットが 0.48 とし、慣性項は用いなかった。また、シグモイド関数は値域を -0.5 から 0.5 とし、状態評価値は、このうち -0.4 から 0.4 の部分を 0 から 1 に線形変換し、動作は、やはり -0.4 から 0.4 の部分を -3.0 から 3.0 に線形変換して用いた。

学習曲線を Fig. 8 に示す。位置、速度の初期値をそれぞれ 0.25 、 1.33 間隔の格子上の点から、物理的に登ることができない点を除いた 37 点について、ゴール達成まで

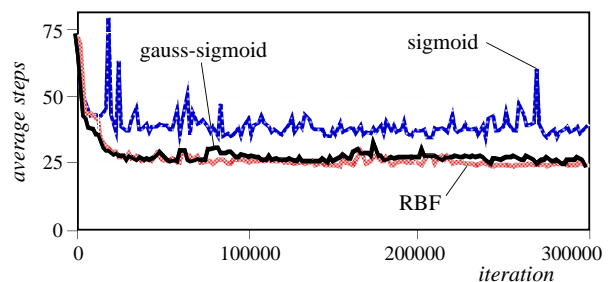


Fig. 8 Comparison of the learning curve in the hill-car task.

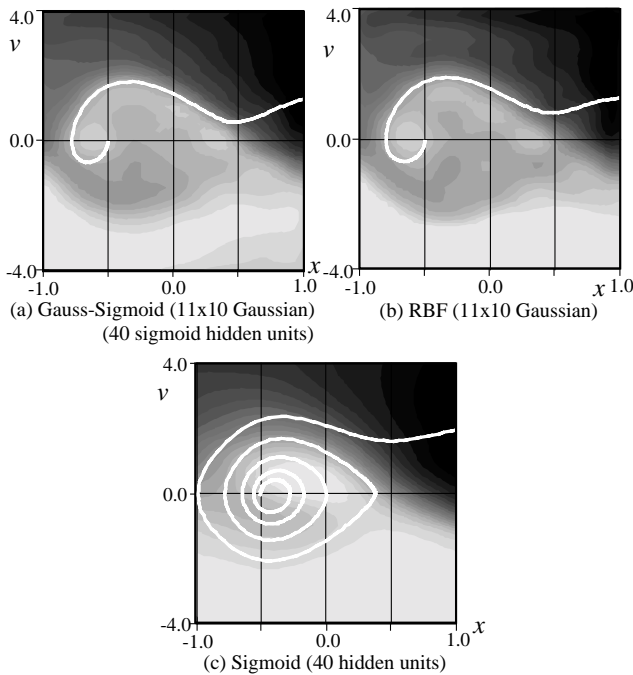


Fig. 9 Value function and locus of the car in hill-car task. The initial state of the locus is $(x, v) = (-0.5, 0.0)$ that is the bottom of the slope.

の平均所要ステップ数を 2000 回の学習ごとにプロットした。これより、Gauss-Sigmoid NN は、シグモイドユニットを用いているにもかかわらず、RBF とほぼ同等の学習速度を実現していることがわかる。また、Gauss-Sigmoid NN とパラメータ数をそろえるため、シグモイド NN の中間層ニューロン数を増大させてシミュレーションを行ったが、学習結果に大きな変化はなかった。

また、3つのネットワークのそれぞれについて、300000 回学習後の状態評価値の分布と $(x, v) = (-0.5, 0.0)$ をスタートとした場合の車の軌跡を Fig. 9 に示す。いずれの場合も、位置 x 、速度 v とともに大きい右上の部分では状態評価値が大きくなっており、離れるにしたがって、その値が小さくなっていることがわかる。また、右側に一気に推力を出して登れるか登れないかの境界部分では、状態評価値が急激に変化していることがわかる。また、Gauss-Sigmoid および RBF ネットでは、 $(x, v) = (-0.85, 0.0)$ 付近で状態評価値の明確な尾根を観察でき、 $(x, v) = (-0.5, 0.0)$ からスタートした場合に、最初に左に登った後に右側の斜面を登ることができていることがわかる。一方、シグモイド NN の場合は、明確な尾根が観察できず、右と左の斜面を何回も行ったり来たりしないと登れない事がわかる。また、この時に Gauss-Sigmoid NN を用いた場合の車の位置 x 、速度 v 、推力 $action$ 、および状態評価値 P の時間変化の様子を Fig. 10 に示す。最初は左に力を出して左側の斜面を登り、 $x = -0.78$ 付近まで登ったところで力を右側に切り替え、勢いをつけて右側の斜面を登り続けていることがわかる。状態評価値は本来単調に増加すべきであるが、ここでは、あまり滑らかになっていない。これは、車が登れるか登れないかの境界で状態評価値が急激に変化している近くを車が通っているため、ちょっとした位置や速度の変化が状態評価値の大きな変化をもたらしていると考えられる。

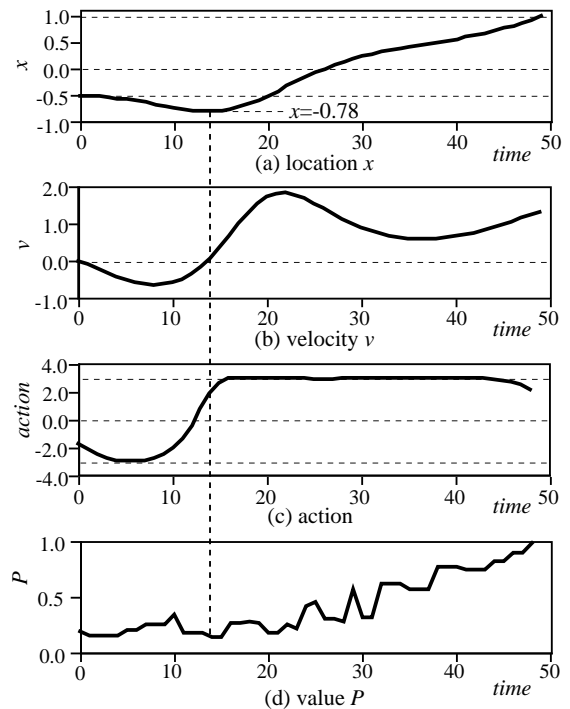


Fig. 10 The change of location x , velocity v , action and the value P of the car when it starts from $(x, v) = (-0.5, 0.0)$.

6. 結論

入力信号をガウシアンユニットで局所化し、その出力をシグモイド型ニューラルネットワークで処理する Gauss-Sigmoid ニューラルネットワークを提案した。これにより、強い非線形関数を近似する場合でも学習が安定するとともに、シグモイド NN の中間層において、大域的な表現も獲得できることを示した。そして、教師あり学習のシミュレーションによってその利点を確認するとともに、強化学習のシミュレーションによって、シグモイド NN を用いているにもかかわらず、入力にガウシアンを挿入することで、より良い解が得られることを示した。

謝辞

本研究の一部は、文部省科研費(#10450165)および日本学術振興会未来開拓学術研究プロジェクト「生物型適応システム」(JSPS-RFTF96I00105)の補助の下で行われた。

Reference

- [1] Boyan, J. A. & Moore, A. W., "Generalization in Reinforcement Learning: safely approximating the value function", in *Advances in Neural Information Processing Systems*, Vol. 7, pp. 369-376, The MIT Press (1995)
- [2] Sutton, R.S., "Generalization in Reinforcement Learning: Successful Examples Using Space Coarse Coding", In *Advanced in Neural Information Processing System*, Vol. 8, pp. 1038-1044 (1996)
- [3] Gordon, G. J., "Stable Function Approximation in Dynamic Programming", *Proc. of the 12th ICML*, pp. 261-268 (1995)
- [4] Shibata, K. & Ito, K., "Gauss-Sigmoid Neural Network", *Proc. of IJCNN'99*, #747 (1999)
- [5] 前原伸一, 杉坂政典, 伊藤宏司, 柴田克成, "Gauss-Sigmoid ニューラルネットワークを用いた強化学習の安定性", 計測自動制御学会九州支部第 19 回学術講演会予稿集, pp. 475-478 (2000)