“メモリー”付きニューロン（慣性系ニューロン）を用いた畳み込み積分型勾配法のニューラルネットワーク実現*

堀江 亮太**・相吉英太郎**・石井 秀教**

The Neural Networks Realization of a Gradient Method with a Convolution Integral by Using Neurons with “Memory”(Inertial Neurons)*

Ryota HORIE**, Eitaro AIYOSHI** and Hidenori ISHII**

First, a new type of models for trajectory methods to solve optimization problems is considered. In this new model, a velocity of the trajectory is given by a convolution integral form with all gradients of the minimization function on the trajectory for the past time. The new trajectory method can be called a gradient method with the optimizer's “memory” with respect to the past gradient information, and the model can be transformed into the second order differential equation model whose trajectory tides over trapping into local optima under a suitable initial velocity.

Next, in order to solve quadratic programming problems with variables constrained on the closed interval $[0,1]$'s, the gradient method with “memory” is realized by neural networks as operational circuits composed of neurons, each of which has two integral elements. The trajectory by realized neural networks has possibility to overcome trapping into local minima, while the Hopfield type with first order differential equation model traps into them.

Last, the numerical simulation results for simple test problems demonstrate properties of these presented neural networks.

1. 緒論

0-1 組合せ最適化問題をニューラルネットワークで解くことが、参考文献 1),2) の研究で脚光を浴びて以来、ニューロン間の相互結合構造を介した帰還構造を前提とし、ニューロン素子特性に様々な非線形特性や揺らぎを仮定したうえで、ネットワーク全体のダイナミックな状態遷移特性を、計算機パワーによるシミュレーションで確認するという方法論が展開されてきた。これに対して筆者の一人は、多重 1 次補間とペナルティ法に基づいて行われる組合せ最適化問題の連続的最適化問題への変換について、まずニューラルネットとは独立にその等価性を考察したうえで、その連続問題をどのようなダイナミクスモデルで解くかを明示し、その実現回路としてニューラルネットワークの構造とニューロン素子特性を演繹的に導入すべきことを、参考文献 3) で指摘した。そして連

続的最適化問題を連続的ダイナミクスで解く場合、局所的最適解への停留をある程度回避できる高階微分方程式最小化モデルと、それを実現した多層帰還型ニューラルネットワークを、参考文献 4) で導出した。このニューラルネットワークは多層構造という点で構造的には新鮮さがあるものの、ニューロン素子自体は単一の積分器と、線形または非線形の出力特性を有するもので、Hopfield 型ニューラルネットワークのニューロン素子と本質的には同質であるといえる。

本論文では、最小化モデル自体は 1 階微分方程式に基づく勾配系最小化モデルでありながら、速度方向を過去の勾配情報の畳み込み積分により生成するモデルを想定する。このような畳み込み積分型勾配法モデルは、最適化演算器がいわば“記憶機能”を有しているものといえ、このモデルをニューラルネットワークで実現すると、構造的には相互結合構造でありながら、それを構成するニューロン素子が 2 個の積分器を内蔵する慣性系モデル⁵⁾となる。このようなニューラルネットワークを各変数が閉区間 $[0,1]$ 上に制約された、超単位立方体上の 2 次関数や多重 1 次関数の簡単な最小化問題に対して適用して、その

* 原稿受付 1996 年 10 月 18 日

** 慶應義塾大学 理工学部 Faculty of Science and Technology, Keio University; 3-14-1 Hiyoshi, Kouhoku ward, Yokohama city, Kanagawa 223-0061, JAPAN

Key Words: optimization problem, gradient method, convolution integral, neural network.

最小化の収束特性を示す。とくに、Hopfield 型ニューラルネットワークでは、本質的に克服できない局所的最適解への引き込み現象の回避が、本モデルによって可能になることを確認する。なお、参考文献7)では、散逸項を非線形化した2階微分方程式モデルを直ちにニューラルネットワーク全体の内部状態モデルとして与えているが、このようないわゆるカオス的な振動現象も、このニューロン素子の慣性系のダイナミックスをさらに非線形化することによって実現可能であることも強調しておく。

2. 畳み込み積分型勾配法モデル

まず無制約最小化問題

$$\min_{\mathbf{x}} E(\mathbf{x}) \quad (1)$$

に対する勾配法のモデルとして、過去の勾配情報の畳み込み積分によって、速度方向を生成するモデル

$$\frac{dx_i(t)}{dt} = -c \int_0^t e^{-a(t-\tau)} \frac{\partial E(\mathbf{x}(\tau))}{\partial x_i} d\tau \quad i=1, \dots, n \quad (2)$$

を考える。ただし $\mathbf{x} \in \mathbf{R}^n$ とし、関数 $E: \mathbf{R}^n \rightarrow \mathbf{R}^1$ は連続微分可能であるとする。ここで右辺の指数関数は、いわゆる線形システムの時刻 τ における時定数 $1/a$ のインパルス応答を表す重み関数である。また $c > 0$ とする。このモデルのダイナミックスは、単純な最急降下法のようにその時点の勾配の逆方向に状態遷移するのではなく、過去の軌道上の勾配の影響を受けて状態遷移する。ただし重み関数 $e^{-a(t-\tau)}$ により、過去の勾配情報の影響は、過去にさかのぼるほど指数関数的に低減する。いわば過去の勾配に関して“記憶”を持つため、勾配がゼロの点においても、過去の勾配の影響により状態遷移して、この結果、通常最急降下法の最小化モデルによる軌道が停留する局所的最小点を乗り切る能力があるものと期待される。

ここで(2)式の両辺を微分すると

$$\begin{aligned} \frac{d^2 x_i(t)}{dt^2} &= c \left[a e^{-at} \int_0^t e^{a\tau} \frac{\partial E(\mathbf{x}(\tau))}{\partial x_i} d\tau - \frac{\partial E(\mathbf{x}(t))}{\partial x_i} \right] \\ &= c \left[a \int_0^t e^{-a(t-\tau)} \frac{\partial E(\mathbf{x}(\tau))}{\partial x_i} d\tau - \frac{\partial E(\mathbf{x}(t))}{\partial x_i} \right] \\ &= -a \frac{dx_i(t)}{dt} - c \frac{\partial E(\mathbf{x}(t))}{\partial x_i} \end{aligned} \quad (3)$$

となる。したがって、(2)式は2階微分方程式

$$\frac{d^2 x_i(t)}{dt^2} + a \frac{dx_i(t)}{dt} = -c \frac{\partial E(\mathbf{x}(t))}{\partial x_i} \quad i=1, \dots, n \quad (4)$$

からなるモデルと等価なダイナミックスを持つことがわかる。

3. ニューラルネットワーク実現

この畳み込み積分型勾配法モデルを上下限制約付き問題

$$\begin{aligned} \min_{\mathbf{x}} E(\mathbf{x}) \\ \text{subj.to } 0 \leq x_i \leq 1, \quad i=1, \dots, n \end{aligned} \quad (5)$$

に対して適用するには、参考文献3)のHopfield型モデルと同様に、上下限への近さに応じたブレーキ係数を乗じたモデル

$$\begin{aligned} \frac{dx_i(t)}{dt} &= -cx_i(t)\{1-x_i(t)\} \\ &\times \left(\int_0^t e^{-a(t-\tau)} \frac{\partial E(\mathbf{x}(\tau))}{\partial x_i} d\tau \right) \quad i=1, \dots, n \end{aligned} \quad (6)$$

を考えることができる。

この微分方程式で記述されたモデルをアナログ演算器で実現するために、出力量 x_i と区別した状態量 u_i を用い、出力方程式を

$$x_i = f(u_i), \quad i=1, \dots, n \quad (7)$$

とした素子を考える。(7)式で $x_i = x_i(t)$, $u_i = u_i(t)$ とおいて、時間 t で全微分すると、

$$\frac{dx_i(t)}{dt} = \frac{dx_i(t)}{du_i} \frac{du_i(t)}{dt}, \quad i=1, \dots, n \quad (8)$$

となるので、(6)式の実現のためには、

$$\frac{df(u_i(t))}{du_i} = x_i(t)\{1-x_i(t)\}, \quad i=1, \dots, n \quad (9)$$

$$\frac{du_i(t)}{dt} = -c \int_0^t e^{-a(t-\tau)} \frac{\partial E(\mathbf{x}(\tau))}{\partial x_i} d\tau \quad i=1, \dots, n \quad (10)$$

とすればよいことがわかる。(9)式からシグモイド関数

$$x_i = f(u_i) = \frac{1}{1 + \exp(-u_i)}, \quad i=1, \dots, n \quad (11)$$

が出力方程式として得られることは、参考文献3)で指摘したとおりである。また(10)式からは、(2)式から(4)式が得られるのと同様に、

$$\frac{d^2 u_i(t)}{dt^2} + a \frac{du_i(t)}{dt} = -c \frac{\partial E(\mathbf{x}(t))}{\partial x_i} \quad i=1, \dots, n \quad (12)$$

が得られる。

ここで, u_i を第 i ニューロンの内部状態, x_i をその出力量にとり, 最小化関数 E として, 2 次関数

$$E(x) = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n w_{ij} x_i x_j + \sum_{i=1}^n \theta_i x_i \quad (13)$$

にとると,

$$\frac{\partial E(x)}{\partial x_i} = \sum_{j=1}^n w_{ij} x_j + \theta_i \quad i = 1, \dots, n \quad (14)$$

であるから, Fig. 1 のようなネットワーク構造と, Fig. 2 のようなニューロン素子の内部構造が得られる. ニューロン素子の内部構造は積分器 2 個を内蔵し, その内部状態自体が 2 階微分方程式系として振る舞うことに注意されたい. そして定数倍 ($-a$) を介した帰還路付きの左側の積分器が, 過去の勾配情報に関する記憶機能を実現していることになる. なお, Fig. 2 の線形散逸項に相当するニューロン内の $\dot{u}_i$ のフィードバック部分を定数 a 倍から, 一般の非線形入出力関数 $f(\dot{u}_i, t)$ とすると, Fig. 3 のようになり, 参考文献 7) のカオスの振動現象を内部状態として発現するモデルが実現される.

本論文で提案したニューラルネットワークは, このように内部状態が 2 階微分方程式モデルとなるが, 最小化変

Fig. 1 Structure of neural networks

Fig. 2 Inner structure of neurons

Fig. 3 Inner structure of neurons with a nonlinear dissipative term

数が 2 階微分方程式モデルとなるニューラルネットワークとして, 参考文献 4) のモデルの高階微分方程式最小化モデルにおいて, $M=2$, $a_1=1$, $a_2=1/a$ として得られるモデルがある. この参考文献 4) のモデルと本論文のモデルは, $a=c$ のときにある変数の対応関係のもとで等価になるが, このことの詳細については付録 1. に譲る.

4. シミュレーション結果

提案したニューラルネットワークの最小化特性を, 計算機シミュレーションによって調べた. 以下では, 時間刻み幅を $\Delta t=0.01$ とした Runge-Kutta 法でシミュレーションを行った. またモデルパラメータはつねに $c=1.0$ とした.

4.1 ニューロン素子の動特性

まず, 基本となる 1 個のニューロン素子の基本動特性を, 1 変数最小化問題を

$$\begin{aligned} \min_x & -x^2 + 1.4x \\ \text{subj.to } & 0 \leq x \leq 1 \end{aligned} \quad (15)$$

で示す. この問題の最小化関数の形状は Fig. 4 である.

Fig. 4 Form of the one dimensional minimization function

最小化変数の初期値 $(x(0), dx(0)/dt)$ を $0.1 \leq x(0) \leq 0.9$, $-1.0 \leq dx(0)/dt \leq 1.0$ の領域の $5 \times 11 = 55$ 個の格子点に取り, Runge-Kutta 法を時刻 $T = 20.0$ まで実行した. また $a = 1.0$ とした. ただし, ニューラルネットワークの初期内部状態 $(u(0), du(0)/dt)$ は, x の初期状態 $(x(0), dx(0)/dt)$ を用いて, (8), (9), (11) 式より逆算した式

$$\begin{cases} u(0) = -\ln \left| \frac{1}{x(0)} - 1 \right| \\ \frac{du(0)}{dt} = \frac{1}{x(0)(1-x(0))} \frac{dx(0)}{dt} \end{cases} \quad (16)$$

によって求めた. 最小化変数の状態空間における軌道 $(x(t), dx(t)/dt)$ を Fig. 5(a) に, そのときの内部状態の状態空間軌道 $(u(t), du(t)/dt)$ を Fig. 5(b) に示す. Fig. 5(a) より, 極大点 $x = 0.7$ の両側からでも適当な初速度を持たせることで, この点を乗り越えて状態遷移することができ, 上下制限の境界上の局所的最適解 $x = 0.0$ もしくは $x = 1.0$ に収束する. また, Fig. 5(b) より内部状態変数の速度 $du(t)/dt$ は 0 でない一定の値に収束し, $u(t)$ の大きさは増加し続ける.

なお, 付録 1 に示す参考文献 4) のモデルのニューラルネットワークを, この 1 変数問題に対して, 最小化変数の同じ初期状態 $(X(0), \dot{X}(0)) = (x(0), \dot{x}(0))$ から $a = 1.0$ として数値積分した結果を Fig. 6 に示す. ただし, ニューラルネットワークの初期状態はそれぞれ最小化変数の初期状態から逆算して与えた. Fig. 5(a) とは異なる軌道で局所的最小点 0.1 に収束する. また, 同じ初期状態からでも異なる最小点に収束する軌道も確認される. ここで, 本論文のモデルでは最小化変数が閉区間 $[0, 1]$ に制約されているのに対応して, 参考文献 4) のモデルでは

$$0 \leq \frac{1}{a} \frac{dX(t)}{dt} + X(t) \leq 1 \quad (17)$$

となり, 初期状態もこの関係を満たしていなければならず, $X(0)$ の値によっては, 任意の初速度 $\dot{X}(0)$ はとれない欠点がある. 逆に, 初期位置が $0 < X(0) < 1$ を満たさなくても (17) 式を満たすように初速度を与えれば, 最小化変数の制約領域外からでも, 制約領域内の最小点を探索できるモデルである. なお, Fig. 6 の 2 本の斜線の囲む帯状領域内側が (17) 式を満たす領域である.

4.2 2 変数最小化問題

参考文献 4), 6) の 2 変数多重 1 次関数最小化問題

$$\begin{aligned} \min_{\mathbf{x}} \quad & -\frac{3}{2}x_1x_2 + x_1 + x_2 \\ \text{subj.to} \quad & 0 \leq x_i \leq 1, \quad i = 1, 2 \end{aligned} \quad (18)$$

に対して, Hopfield 型ニューラルネットワークとの比較結

(a) Trajectories in the phase-space of minimization variables

(b) Trajectories in the phase-space of inner state variables

Fig. 5 Dynamical property of a neuron

果を示す. この最小化問題は Fig. 7 に示すように, 大域的最低点を上下制限の境界上の $(0, 0)$ に持ち, 最小値は 0 である. また, 局所的最低点を $(1, 1)$ に持つ. Runge-Kutta 法の終了条件は, 最小化変数の遷移変動幅が 10^{-6} 以下となったステップが 5 回続いたとき, あるいはその更新回数が 3 万回を越えたときとした. また, ニューラルネットワークの初期内部状態 $(u_i(0), du_i(0)/dt), i = 1, 2$ は, 4.1 と同様に $\mathbf{x}$ に関する初期状態 $(x_i(0), dx_i(0)/dt), i = 1, 2$ を用いて, 各成分ごとに (16) 式より逆算して指定した. 初期点 $\mathbf{x}(0)$ に対して, 適切に初速度 $d\mathbf{x}(0)/dt$ が

Fig. 6 Trajectories in the phase-space of minimization variables by second order differential equation in the ref.4)

Fig. 7 Form of the two dimensional minimization function

与えられれば、大域的最小解への初期点引き込み領域が拡大することを確認するために、最小化変数の初期点を $0.1 \leq x_i \leq 0.9, i=1,2$ の領域内の $9 \times 9 = 81$ 個の 0.1 刻みの格子状にとり、初速度方向 $d\mathbf{x}(0)/dt$ を大域的最小解の方向に設定し、その大きさ $\|\dot{\mathbf{x}}(0)\|$ のみを変化させてシミュレーションを行った。係数 a は $a=1.0$ とした。また比較のため、同じ初期点からの Hopfield 型ニューラルネットワークの最小化変数の軌道群を Fig. 8(a) に、提案した“メモリー”付きニューラルネットワークで、初速度の大きさを 0 としたときの最小化変数の軌道群を Fig. 8(b) に、また初速度の大きさを 1.0 としたときを Fig. 8(c) に

(a) Hopfield type

(b) Type with the memory ($\|\dot{\mathbf{x}}\| = 0$)

(c) Type with the memory ($\|\dot{\mathbf{x}}\| = 1.0$)

Fig. 8 Trajectories of minimization variables in the two dimensional problems

示す. ニューロンの内部状態に関して2階微分方程式のダイナミクスを有する本ニューラルネットワークでも, 初速度の大きさが0であれば, 状態遷移軌道は引き込み領域内で局所的最小解へ向けて加速するだけで, 従来のHopfield型のニューラルネットワークの結果とほとんど変わらない. これに対して, 初速度の大きさを十分に与えると, 多重1次関数が形成する“尾根”を乗り越え, すべて大域的最小解へ収束していることがFig. 8(c)よりわかり, 引き込み領域の拡大が確認される.

つぎに, 上と同じ条件のもとで, 大域的最適解への収束率に対する, 初速度の大きさ $\|\dot{\mathbf{x}}(0)\|$ の影響および係数 α の影響をより詳しく調べた. Fig. 9に初速度の大きさの影響を示す. このとき係数 α は一定値 $\alpha=1.0$ とした. 明らかに初速度が大きくなるほど収束率はよくなることが確認される. 上下制限約の境界に軌道が近づくほど, ブレーキ効果が生じるため, 2階のダイナミクスを用いた一般の無制約最小化のように, 勢いをつけすぎて大域的最小解を通りすぎるようなことは起こらない.

Fig. 9 Variation of convergence ratio with norm of initial velocities

つぎに, Fig. 10に係数 α に対する影響を示す. このとき初速度の大きさは $\|\dot{\mathbf{x}}(0)\|=0.1$ と一定にした. この係数 α は, ニューロンの内部状態のダイナミクスを表す(12)式の制動係数に相当する. したがって, この大きさが大きいとき, 慣性力より制動力が相対的に大きくなり, この最小化特性是最急降下法に近付き, 逆に小さいとき, 慣性力の方が相対的に大きくなって, 初速度の効果がより強く効いてきて, その軌道は“軟らかな”動きをするようになり, 大域的最小解への収束率が向上する.

4.3 25変数最小化問題

参考文献6)と同じであり, 参考文献4)でも用いられた, ランダムに選んだ6種類の係数を有する25変数多重1次関数最小化問題に対して, Runge-Kutta法でシミュレーションを行った. 終了条件は, 4.2の2変数問題と同

Fig. 10 Variation of convergence ratio with parameter α

じである. 最小化変数の初期点を25次元の超単位立方体内 $0.001 \leq x_i \leq 0.999, i=1, \dots, 25$ の範囲の乱数で100通り与えた. ただし $\alpha=1.0$ と固定した. このとき,

- (i) Hopfield型ニューラルネットワーク
- (ii) 初速度方向を大域的最小解方向に設定し, 大きさを, $\|\dot{\mathbf{x}}(0)\|=1.0$ とした場合の提案ニューラルネットワーク
- (iii) 初速度方向を大域的最小解方向に設定し, 大きさを, $\|\dot{\mathbf{x}}(0)\|=5.0$ とした場合の提案ニューラルネットワーク
- (iv) 初速度の各成分を $-1 \leq \dot{x}_i \leq 1, i=1, \dots, 25$ の範囲の乱数で与えた提案ニューラルネットワーク

の4通りについて, 各100通りの初期点 $\mathbf{x}(0)$ からの大域的最小解への収束率を示したのがTable 1である. 初速度を初期点から大域的最小解へ方向に設定した場合には, 提案ニューラルネットワークはHopfield型ニューラルネットワークより大域的最小解への収束率が改善され, また初速度の大きさを大きくするほど収束率が高くなることがわかる. なお初速度の各成分を乱数で与えたときは, 平均化されるため必ずしも収束率が改善されないことは明らかである.

一般に, 閉区間 $[0,1]$ 内に制約された2次関数最小化問題では, 2次関数が非凸関数であれば制約領域とその関

Table 1 Convergence ratio of the problems with 25 variables

problem	(i)	(ii)	(iii)	(iv)
1	9%	19%	100%	1%
2	48%	69%	100%	13%
3	49%	93%	100%	33%
4	56%	86%	100%	28%
5	90%	100%	100%	53%
6	50%	64%	100%	75%

数の位置関係により複数の局所的な最小解が存在する。これらの最小解を軌道法で探索するとき、1階常微分方程式で記述される最急降下法とよばれる勾配法モデルの探索軌道は、初期位置によってどの最小解に収束するかが決まり、ある最小解に収束する初期位置領域は、その最小解の「引き込み領域」と呼ばれる。Hopfield型ニューラルネットワークは1階常微分方程式で記述される最急降下法に基づくため、その引き込み領域は初期位置のみで決定される。これに対して2階常微分方程式で記述される慣性系モデルでは、初期位置だけでなく初速度の与え方によっても、どの最小解に収束するかが決まり、同じ初期位置からの引き込み領域であっても初速度によってその領域が変化する自由度を有している。Table 1の結果はこのことを示している。すなわち (iii) のように、適切な初速度が与えられれば、大域的な最小解への引き込み領域は全領域に拡大し、逆にこのような自由度があるために、(iv) のように無作為に初速度を与えては、問題によってはHopfield型ニューラルネットワークより収束率が悪化する。

5. 結論と今後の課題

無制約最小化問題に対する1階微分方程式の最小化モデルにおいて、速度方向を過去の勾配情報の畳み込み積分により生成する、畳み込み積分型勾配法を提案した。このモデルは2階微分方程式モデルと等価であるが、このモデルを上下制限付き2次関数最小化問題に適用した場合、過去の勾配情報に関するメモリーとしての役割を有する慣性項の存在する2階微分方程式モデルによって、ニューロンの内部状態が記述されることを明らかにした。この機能によって、ニューロンの内部状態の初速度を適切に設定すれば、ニューロンからの出力変数である最小化変数軌道の局所的な最小点への引き込み領域を、大幅に拡大させることができることが確認された。なお、参考文献6)においても、上下制限付き境界付近でのプレーキの掛けかたを急激にすることで、大域的な最適解への引き込み領域を拡大できることが指摘されているが、このモデルの本質は1階微分方程式モデルであることや、そのニューラルネットワーク実現が容易でないことなどが指摘される。

ところで、シミュレーション結果が示すとおり、ランダムに初期速度を指定しては、逆に大域的な最小点への収束性は統計的に劣化する。したがって、与えられた初期点に対して大域的な最小解へ収束する適切な初速度を選択する、という課題が残されている。これに対しては、本ニューラルネットワークを繰り返して動作させることによって、メタレベルにおいてたとえば遺伝的アルゴリズム

ムなどの試行探索法⁸⁾を用いて、このような適切な初速度を探索することが考えられる。とくに、生体生理学的に言えば、内部状態量に関する情報は外部から付加不可能な量であるがために、ニューロンの最良の初期内部状態やその初期速度決定問題には、多数のニューラルネットワーク個体を用いた自律的進化型アルゴリズムが適するものと思われるが、このことに関する具体的な考察は、新しいニューロンモデルに基づく新最適化ダイナミックスの原理的提案という本論文の主旨に必ずしもそぐわないため、別の機会に譲ることとする。

参考文献

- 1) J. J. Hopfield and D. W. Tank: "Neural" computation of decision in optimization problems; *Biol. Cybern.*, Vol. 52, pp. 141-152 (1985)
- 2) D. W. Tank and J. J. Hopfield: Simple "Neural" optimization networks: An A/D converter, signal decision circuit and linear programming circuit; *IEEE Trans. CAS*, Vol. CAS-33, pp. 533-541 (1986)
- 3) 相吉, 吉川: ニューラルネットワークによる最適化: その疑問と課題—もう一つの接近法—; 計測と制御, Vol. 34, No. 5, pp. 358-366 (1995)
- 4) 相吉, 阿部: 高階微分方程式最小化モデルの多層帰還型ニューラルネットワークによる実現; 電学論C, Vol. 115, No. 10, pp. 1151-1157 (1995)
- 5) S. Incerti, V. Parisi and F. Zirilli: A new method for solving nonlinear simultaneous equations; *SIAM J. Numerical Analysis*, Vol. 16, No. 5, pp. 779-789 (1979)
- 6) 大堀, 山本, 藤, 渡邊: 勾配加速ホップフィールドニューラルネットワークによる組合せ最適化問題の解法; 電学論C, Vol. 114, No. 4, pp. 420-425 (1994)
- 7) 谷: カオスの最急降下法を適用したニューラルネットワークにおける学習および記憶想起の動特性について; 信学論(A), Vol. J74-A, No. 8, pp. 1208-1215 (1991)
- 8) 相吉, 三室: 大域的な最適化のためのメタ最適化問題と遺伝的アルゴリズムによる解法; 計測自動制御学会論文集, Vol. 28, No. 8, pp. 999-1006 (1992)

付 録

付録 1. 参考文献4)のモデルと本論文のモデルの関係
参考文献4)のモデルの高階微分方程式最小化モデルにおいて、 $M=2$, $a_1=1$, $a_2=1/a$ とすると

$$\frac{1}{a} \frac{d^2 X_i(t)}{dt^2} + \frac{dX_i(t)}{dt} = - \left\{ \frac{1}{a} \frac{dX_i(t)}{dt} + X_i(t) \right\} \times \left\{ 1 - \left(\frac{1}{a} \frac{dX_i(t)}{dt} + X_i(t) \right) \right\} \frac{\partial E(X(t))}{\partial X_i} \quad i=1, \dots, n \quad (A1)$$

となる。これを実現した2層帰還型ニューラルネットワークは、各層のニューロンの内部状態を (U_{n+i}, U_i) , 出力を (Y_{n+i}, Y_i) とおくと

$$\frac{dU_{n+i}(t)}{dt} = -a \sum_{j=1}^n w_{ij} Y_j(t) - \theta_i \quad (\text{A2a})$$

$$\frac{dU_i(t)}{dt} = Y_{n+i}(t) - aY_i(t) \quad (\text{A2b})$$

$$Y_{n+i}(t) = \frac{1}{1 + \exp(-U_{n+i}(t))} \quad (\text{A2c})$$

$$Y_i(t) = U_i(t) \quad (\text{A2d})$$

$$i = 1, \dots, n$$

で表現される。ただし、最小化変数と出力変数の関係は

$$Y_i(t) = \frac{1}{a} X_i(t) \quad (\text{A3a})$$

$$Y_{n+i}(t) = \frac{1}{a} \frac{dX_i(t)}{dt} + X_i(t) \quad (\text{A3b})$$

$$i = 1, \dots, n$$

である。なお、参考文献 4) のモデルの変数は区別して大文字で記す。ここで、(A2b) に (A2d) を代入すると、

$$Y_{n+i}(t) = \frac{dY_i(t)}{dt} + aY_i(t), \quad i = 1, \dots, n \quad (\text{A4})$$

となる。また、(A2a) の両辺を時間微分したものに (A2a) に a を乗じたものを加えて、(A4) を用いると

$$\begin{aligned} & \frac{d^2 U_{n+i}(t)}{dt^2} + a \frac{dU_{n+i}(t)}{dt} \\ &= -a \left(\sum_{j=1}^n w_{ij} Y_{n+i}(t) + \theta_i \right), \quad i = 1, \dots, n \end{aligned} \quad (\text{A5})$$

となる。ここで、参考文献 4) のモデルと本論文のモデルとで変数の対応を Table A1 のようにおくと、 $a=c$ のときに限って、(A5) は (12) 式と一致することがわかる。すなわち、本論文のモデルの最小化変数 $\mathbf{x}$ と参考文献 4) のモデルの最小化変数 $\mathbf{X}$ の軌道はまったく異なるが、Table A1 の変数の対応関係のもとで等価であるといえる。なお、 $a \neq c$ ならば、この対応関係のもとでも等価でないことに注意されたい。

Table A1 Correspondence between the presented model and the ref. 4)'s model

the presented model	the ref. 4)'s model
$u_i(t)$	$U_{n+i}(t)$
$\dot{u}_i(t)$	$-a \sum_{j=1}^n w_{ij} Y_j(t) - \theta_i$
$x_i(t)$	$Y_{n+i}(t)$ $= \frac{1}{a} \frac{dX_i(t)}{dt} + X_i(t)$