神経回路網におけるカオスダイナミクスの制御†

酒井正夫*・本間経康**・阿部健一*

Complexity Control Methods of Chaos Dynamics in Recurrent Neural Networks†

Masao SAKAI*, Noriyasu HONMA** and Kenichi ABE*

This paper demonstrates that the Lyapunov exponents of recurrent neural networks can be controlled by our proposed methods. One of the control methods minimizes a squared error $e_\lambda = (\lambda - \lambda^{obj})^2/2$ by a gradient method, where λ is the largest Lyapunov exponent of the network and λ^{obj} is a desired exponent. λ implying the dynamical complexity is calculated by observing the state transition for a long-term period. This method is, however, computationally expensive for large-scale recurrent networks and the control is unstable for recurrent networks with chaotic dynamics since a gradient collection through time diverges due to the chaotic instability. We also propose an approximation method in order to reduce the computational cost and realize a “stable” control for chaotic networks. The new method is based on a stochastic relation which allows us to calculate the collection through time in a fashion without time evolution. Simulation results show that the approximation method can control the exponent for recurrent networks with chaotic dynamics under a restriction.

Key Words: recurrent neural networks, complexity, chaos

1. はじめに

フィードバック結合をもつリカレントニューラルネットワーク (Recurrent Neural Networks: 以下 RNN と略す) は、構成ニューロン間の自由度の高い相互作用が生み出す複雑なダイナミクスを利用した動的システムの同定問題、制御問題などへの応用が期待されている¹⁾。また、Hopfield らによって階層型ニューラルネットワーク (Feedforward Neural Networks: 以下 FNN と略す) とは異なる認識機構 (最適化機構) をもつ完全結合型 RNN²⁾が提案されており、ノイズ画像の優れた復元能力などが報告されている³⁾。この場合、RNN のシナプス結合の値が制限されているため、そのダイナミクスは一点へ向かう (ポイントアトラクタ) という意味で非常に単純である。しかし、Freeman らの動物実験の結果によると、生物の認識機構には、カオスダイナミクスおよびカオスの辺縁 (the edge of chaos) のダイナミクスが深く関わっていることが報告されている⁴⁾。したがって、RNN を用いて生物のような柔軟な認識機構を実現するためには、ネットワークダイナミクスの複雑さを考慮することが重要な一つの

方法になりうる。

RNN に対する学習法は、BPTT 法、RTRL 法のアルゴリズム⁵⁾を基にしたものが多いが、ネットワークダイナミクスがカオスの場合、学習に必要な微係数の計算がカオスの初期値敏感性⁶⁾により発散してしまうという問題がある。Principe らは、BPTT 法での評価関数に最大リアプノフ指数⁶⁾(ダイナミクスの複雑さ)を考慮した減衰関数を導入することで、微係数の計算の発散をふせぐ学習法を提案している⁷⁾。また、Deco らは写像関数の近似能力に優れる Jordan 型 RNN⁸⁾に対して、減衰関数に加えて効率的に学習するための教師データを決定する “sample-by-sample” 法を導入することで、ダイナミクスの複雑さを考慮した学習が可能であると報告している⁹⁾。しかし、これらの方法は対象となる時系列データを学習するといういわば間接的な手法であり、RNN のダイナミクスの複雑さを直接的に制御するための定式化はなされていない。

著者らはこれまでに、降下法を基に最大リアプノフ指数制御のための定式化を行い、完全結合型 RNN と Jordan 型 RNN のダイナミクスの複雑さを制御する方法を提案している¹⁰⁾。また、この手法と従来の学習法を組み合わせることにより、最大リアプノフ指数に関しても、対象ダイナミクスの複雑さに、より近いモデルが構築できることを示している^{10), 11)}。しかし、この制御法も微係数計算に BPTT 法、RTRL 法のアルゴリズムを用いるため、RNN がカオスダイナミクスをもつ場合にはその精度が十分であるとはいえず、また、ネットワークの構成ニューロン数が増えると莫大な計

† 第 38 回計測自動制御学会学術講演会で発表 (1999・7)

* 東北大学大学院工学研究科 仙台市青葉区荒巻字青葉

** 東北大学医療技術短期大学部 仙台市青葉区星陵町 2-1

* Graduate School of Engineering, Tohoku University, Aoba-ku, Sendai

** College of Medical Sciences, Tohoku University, Aoba-ku, Sendai

(Received June 19, 2000)

(Revised December 21, 2000)

算コストを必要とした。

平澤らによる、一般化学習ネットワーク理論を基にしたダイナミクスの複雑さ制御法では、降下法にランダムサーチを組み合わせることでカオスダイナミクスの制御に成功しているが¹²⁾、必要な計算量はむしろ増加している。

本稿では、構成ニューロン数の多い完全結合型 RNN のダイナミクスの複雑さとネットワークパラメータの関係を統計的に解析して、ダイナミクスの複雑さを制御するための近似的な定式化を行う。シミュレーション結果により、この定式化を用いた近似計算法がこれまでの手法に比べ、構成ニューロン数の多い完全結合型 RNN に対する高い計算コストを低減し、RNN がカオスダイナミクスをもつ場合でも精度の高い複雑さ制御が可能であることを示す。

2. RNN のダイナミクスの複雑さ

ダイナミクスの複雑さを定量化する指標は、種々提案されているが、本稿では、計算コストなども考慮して、最大リアプノフ指数を用いることにする¹⁰⁾。

連立差分方程式

$$\mathbf{x}(t+1) = \mathbf{f}(\mathbf{x}(t)), \quad (1)$$

$$\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_N(t)]^T,$$

$$\mathbf{f}(\mathbf{x}(t)) = [f_1(\mathbf{x}(t)), f_2(\mathbf{x}(t)), \dots, f_N(\mathbf{x}(t))]^T,$$

の最大リアプノフ指数 λ は、次式で計算される¹³⁾。

$$\lambda = \lim_{T \rightarrow \infty} \frac{1}{T} \log \left[\frac{L(T)}{L(0)} \right]. \quad (2)$$

ただし、 $L(t) = \|\mathbf{L}(t)\|$ とし、 $\mathbf{L}(0)$ は $\mathbf{x}(0)$ を中心とした無限小の N 次元球体の直径ベクトル、 $\mathbf{L}(t)$ はその球を (1) 式で t 回遷移させたときに得られる $\mathbf{x}(t)$ を中心とした楕円体の最大の主軸ベクトルである。 $\mathbf{L}(t)$ の計算にアトラクタ近傍でのヤコビ行列での線形化と正規化を用いることで、次式が得られる¹⁰⁾。

$$\lambda = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} \log L(t+1). \quad (3)$$

本稿では、各ニューロンの写像関数にシグモイド関数を用い、 i 番目 ($i = 1, 2, \dots, N$) のニューロンが次式で時間発展してゆく完全結合型 RNN を考える。

$$x_i(t+1) = \frac{1}{1 + \exp(-2s_i(t+1)/s_0)}, \quad (4)$$

$$s_i(t+1) = \sum_{j=1}^N w_{ij} x_j(t) + \theta_i, \quad (t = 1, 2, \dots). \quad (5)$$

ここで、 x_i , s_i はニューロン i の出力、入力で、 N , w_{ij} , s_0 , θ_i は、それぞれニューロン数、ニューロン j との結合重み係数、シグモイド関数の傾き、および閾値である。

(4), (5) 式の RNN の場合、 $\mathbf{L}(t+1) = [L_1(t+1), L_2(t+1), \dots, L_N(t+1)]^T$ の要素は、次式で表される¹⁰⁾。

$$L_i(t+1) = \frac{2}{s_0} X_i(t+1) \sum_{j=1}^N w_{ij} \delta L_j(t). \quad (6)$$

ただし、 $\delta \mathbf{L}(t) = [\delta L_1(t), \delta L_2(t), \dots, \delta L_N(t)]^T$ は

$$\delta \mathbf{L}(t) = \frac{\mathbf{L}(t)}{L(t)}, \quad (7)$$

として、 $X_i(t) = x_i(t)(1 - x_i(t))$ である。よって

$$L(t+1) = \sqrt{\sum_{i=1}^N \{L_i(t+1)\}^2}, \quad (8)$$

であるから、 $\|\delta \mathbf{L}(0)\| = 1$ となる任意の N 次元ベクトルを準備すれば、(3) 式の最大リアプノフ指数は、RNN の出力時系列 $\mathbf{x}(t)$ ($t = 1, 2, \dots, T$) と、結合重み係数 w_{ij} により計算できる。

3. 最大リアプノフ指数の制御

3.1 降下法による制御

ここでは、あるネットワークの最大リアプノフ指数を所望の値に設定する問題を考える。最大リアプノフ指数は、前節で示したように結合重み係数 w_{ij} といったネットワークパラメータによって計算される。したがって、ネットワークダイナミクスが目標とする最大リアプノフ指数 λ^{obj} の複雑さをもつように設定するには、最適なネットワークパラメータ値を設定する問題に帰着される。著者らは、これまでに降下法を用いる設計手法を提案している¹⁰⁾。

いま、ネットワークの最大リアプノフ指数を λ とすると、目標の λ^{obj} に近づけるためのパラメータ w_{ij} の改変量 δw_{ij} は、目標値との二乗誤差を $e_\lambda = (\lambda - \lambda^{obj})^2/2$ として

$$\begin{aligned} \delta w_{ij} &= -\eta \frac{\partial e_\lambda}{\partial w_{ij}}, \\ &= -\eta (\lambda - \lambda^{obj}) \frac{\partial \lambda}{\partial w_{ij}}, \end{aligned} \quad (9)$$

で与えられる。ここで、 $\eta (> 0)$ は学習係数である。従来法は、BPTT 法、RTRL 法⁵⁾などを併用することにより、微係数 $\partial \lambda / \partial w_{ij}$ を (3) 式の定義に従って計算する手法である。

3.2 近似計算法

従来の制御法¹⁰⁾は、(9) 式の微係数 $\partial \lambda / \partial w_{ij}$ の計算に、 $\partial \mathbf{x}(t) / \partial w_{ij}$ ($t = 1, 2, \dots, T$) の時系列データを必要とする。しかし、 T が大きく RNN がカオスダイナミクスを持つ場合は、Principe らも指摘しているように^{7), 9)}、この計算がカオスの初期値敏感性⁶⁾により発散して精度が低下してしまう可能性がある。また、微係数の計算に必要な計算量 (必要な乗算回数) は $O(N^5 T)$ 、メモリ量は $O(N^3)$ となり、ニューロン数が増えると莫大な計算コストになってしまう¹⁰⁾。

著者らは、これらの問題を解決するために、完全結合型 RNN のダイナミクスの複雑さと、RNN の構造上のパラメータ入力分散 σ^2 との関係を定性的に考察し、ネットワーク構造にある拘束条件を設けることで、所望のダイナミクスの複雑さをもつ RNN を設計する近似計算法を提案している¹⁴⁾。以下に、その近似計算法を簡単に述べる。

はじめに、 σ^2 を導入する。(5) 式の w_{ij} を期待値 0 の一様

ランダムに設定し、閾値 θ_i を次式で設定する。

$$\theta_i = -\frac{1}{2} \sum_{j=1}^N w_{ij}. \quad (10)$$

これより、閾値をニューロン j からの寄与 s_{ij} に含めて、入力 s_i を $s_i = \sum_j s_{ij}$ で表す。ただし、 $s_{ij} \equiv w_{ij}(x_j - 1/2)$ である。以下、 z の期待値を $E[z]$ 、分散を $\sigma^2(z)$ で表す。

RNN のダイナミクスがカオスの場合、ニューロンの出力 x_j が区間 $[0,1]$ の一様分布であると仮定すると⁶⁾、 x_j と w_{ij} は互いに独立となるから、期待値は $E[s_{ij}] = E[w_{ij}] \cdot E[x_j - 1/2] = 0$ である。したがって、分散は $\sigma^2(s_{ij}) = E[s_{ij}^2]$ より同様に

$$\sigma^2(s_{ij}) = E[w_{ij}^2] \cdot E[(x_j - 1/2)^2], \quad (11)$$

である。このとき、ニューロン数 N が十分多ければ、 w_{ij}^2 の期待値 $E[w_{ij}^2]$ は平均値と一致するので

$$E[w_{ij}^2] = \frac{1}{N^2} \sum_{i,j} w_{ij}^2. \quad (12)$$

また、 $E[(x_j - 1/2)^2] = 1/12$ であるから、ニューロンの入力分散 σ^2 は中心極限定理より次式で求まる^{14), 15)}。

$$\begin{aligned} \sigma^2 &= N \cdot \sigma^2(s_{ij}), \\ &= \frac{1}{12N} \sum_{i,j} w_{ij}^2. \end{aligned} \quad (13)$$

これを用いると、最大リアプノフ指数を改変するための降下法に必要な微係数は次式で計算できる。

$$\frac{\partial \lambda}{\partial w_{ij}} = \frac{\partial \lambda}{\partial \sigma^2} \cdot \frac{\partial \sigma^2}{\partial w_{ij}}. \quad (14)$$

ここで、 $\partial \sigma^2 / \partial w_{ij}$ は (13) 式より次式で計算できる。

$$\frac{\partial \sigma^2}{\partial w_{ij}} = \frac{1}{6N} w_{ij}. \quad (15)$$

このとき、パラメータ σ^2 と最大リアプノフ指数の関係をシミュレーション結果 (Fig. 1 参照) を基に

$$\lambda \approx \begin{cases} -A_1 \exp(-A_2 \sigma^2) & , (\lambda < 0), \\ B_1 \sigma^2 & , (\lambda > 0), \end{cases} \quad (16)$$

のように経験的に近似し $\partial \lambda / \partial \sigma^2$ を求めることで¹⁴⁾、(14)、(15) 式より近似的な微係数 $\partial \lambda / \partial w_{ij}$ を計算できる。ここで、 A_1, A_2, B_1 は正の定数である。この手法は、微係数 $\partial \lambda / \partial w_{ij}$ の計算に、 $\partial x(t) / \partial w_{ij}$ の時系列データを必要とせず、また、計算量も $O(N^2)$ しか必要としないため、ニューロンの個数が多い大規模ネットワークに対する複雑さ制御にも実用的である。しかし、この近似算法は (16) 式の経験的な定式化に基づいているためその精度は十分とはいえなかった。本稿では、 σ^2 と最大リアプノフ指数の定量的な関係を統計的に解析することにより新しい近似算法を提案する。

3.3 統計的解析

はじめに解析を容易にするために、ニューロン数 N が十分多い完全結合型 RNN の場合、最大リアプノフ指数の計算に用いる要素 $x_i(t)$ 、 w_{ij} ($i, j = 1, 2, \dots, N$, $t = 1, 2, \dots, T$)

が互いに独立であると仮定する。その結果、(3)、(6)、(8) 式と大数の法則より次式のような関係が得られる。

$$\begin{aligned} \lambda &\approx E[\log L(t+1)], \\ &\approx \frac{1}{2} \left\{ \log N + \log E[\{L_i(t+1)\}^2] \right\}, \\ &\approx \frac{1}{2} \left\{ \log \frac{4N}{s_0^2} + \log E[\{X_i(t+1)\}^2] \right. \\ &\quad \left. + \log E \left[\left\{ \sum_{j=1}^N w_{ij} \delta L_j(t) \right\}^2 \right] \right\}. \end{aligned} \quad (17)$$

以下に (17) 式を計算するための、 $E \left[\left\{ \sum_{j=1}^N w_{ij} \delta L_j(t) \right\}^2 \right]$ と $E[\{X_i(t+1)\}^2]$ の求め方を順に述べる。

$$(1) \quad E \left[\left\{ \sum_{j=1}^N w_{ij} \delta L_j(t) \right\}^2 \right] \text{ の計算}$$

w_{ij} が期待値 0 の一様ランダムで設定されたとすると、大数の法則より次式を得る。

$$\begin{aligned} E \left[\left\{ \sum_{j=1}^N w_{ij} \delta L_j(t) \right\}^2 \right] &\approx E \left[\sum_{j=1}^N \{w_{ij} \delta L_j(t)\}^2 \right], \\ &\approx N \cdot E[w_{ij}^2] \cdot E[\{\delta L_j(t)\}^2]. \end{aligned} \quad (18)$$

ここで、 $E[w_{ij}^2]$ は、(12)、(13) 式より次式のように求まる。

$$E[w_{ij}^2] = \frac{12}{N} \sigma^2. \quad (19)$$

また、大数の法則と (7) 式の定義より

$$\begin{aligned} E[\{\delta L_j(t)\}^2] &= \frac{1}{N} \sum_{j=1}^N \{\delta L_j(t)\}^2, \\ &= \frac{1}{N}, \end{aligned} \quad (20)$$

のように求まる。したがって、(18) ~ (20) 式を (17) 式に代入することで次式のような関係が得られる。

$$\lambda \approx \frac{1}{2} \log \left(48 \frac{\sigma^2}{s_0^2} E[\{X_i(t+1)\}^2] \right). \quad (21)$$

$$(2) \quad E[\{X_i(t+1)\}^2] \text{ の計算}$$

(4) 式より、入力を s_i 、出力を x_i とした場合のシグモイド関数のベキ級数展開は、次式のように求まる¹⁶⁾。

$$\begin{aligned} x_i &= 1 / \left\{ 1 + \sum_{n=0}^{\infty} \frac{(-1)^n \cdot 2^n}{n!} \cdot \left(\frac{s_i}{s_0} \right)^n \right\}, \\ &= \sum_{n=0}^{\infty} F_1(n) \cdot \left(\frac{s_i}{s_0} \right)^n, \end{aligned} \quad (22)$$

$$F_1(n) \equiv \begin{cases} 1/2, & (n=0), \\ \sum_{k=1}^n \frac{(-1)^{k+1} \cdot 2^{k-1}}{k!} \cdot F_1(n-k), & (\text{otherwise}). \end{cases}$$

また、(22) 式より x_i^m ($m = 2, 3, \dots$) のベキ級数展開は

$$\begin{aligned} x_i^m &= \sum_{n=0}^{\infty} F_m(n) \cdot \left(\frac{s_i}{s_0} \right)^n, \\ F_m(n) &\equiv \sum_{k=0}^n F_{m-1}(k) \cdot F_{m-1}(n-k), \end{aligned} \quad (23)$$

のように求まることから、 $\{X_i(t+1)\}^2$ は次式のように近似できる。

$$\{X_i(t+1)\}^2 \approx \begin{cases} \sum_{n=0}^{M_1} H(n) \cdot \left(\frac{s_i}{s_0}\right)^n, & \left(\left|\frac{s_i}{s_0}\right| < \alpha\right), \\ 0, & (\text{otherwise}), \end{cases}$$

$$H(n) \equiv \{F_2(n) - 2F_3(n) + F_4(n)\}. \quad (24)$$

ここで、 M_1 は適当な自然数であり、 α は正の定数である。

また、正規分布の確率密度関数を $g(z, E[z], \sigma^2(z))$ と定義すると、入力 s_i の確率密度関数 $g(s_i, 0, \sigma^2)$ は次式のようなべき級数で表される。

$$g(s_i, 0, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{s_i^2}{2\sigma^2}\right),$$

$$= \begin{cases} \sum_{n=0}^{M_2} \frac{R(n)}{\sqrt{\sigma^2}} \cdot \left(\frac{s_i}{\sqrt{\sigma^2}}\right)^{2n}, & \left(\left|\frac{s_i}{\sqrt{\sigma^2}}\right| < \beta\right), \\ 0, & (\text{otherwise}), \end{cases}$$

$$R(n) \equiv \frac{(-1)^n}{\sqrt{2\pi} \cdot 2^n \cdot n!}. \quad (25)$$

ここで、 β は正の定数、 M_2 は適当な自然数であり、 $M_1 = 2M_2$ として表現を簡単化すると、 $E[\{X_i(t+1)\}^2]$ は補正係数 γ を導入して、次式で計算できる。

$$E[\{X_i(t+1)\}^2] = \int_{-\infty}^{\infty} \{X_i(t+1)\}^2 \cdot g(s_i, 0, \sigma^2) ds_i,$$

$$= \sum_{\substack{0 \leq n \leq M_2 \\ 0 \leq k \leq n}} \frac{Q(n, k) \cdot t^{2n+1}}{s_0^{2k} \cdot (\sqrt{\sigma^2})^{2(n-k)+1}} + \sum_{\substack{M_2+1 \leq n \leq 2M_2 \\ n-M_2 \leq k \leq M_2}} \frac{Q(n, k) \cdot t^{2n+1}}{s_0^{2k} \cdot (\sqrt{\sigma^2})^{2(n-k)+1}},$$

$$Q(n, k) \equiv \gamma \cdot \frac{2 \cdot H(2k) \cdot R(n-k)}{(2n+1)},$$

$$t \equiv \min\left(\alpha s_0, \beta \sqrt{\sigma^2}\right). \quad (26)$$

したがって、(26)式を(21)式に代入することで、 λ と σ^2 の統計的解析による近似的関係が求まる。

重み係数をランダムに初期化した構成ニューロン数 $N = 100$ の完全結合型 RNN において、 λ とパラメータ σ^2 の関係を近似計算法によって予測し、実際にシミュレーション

Fig. 1 The Lyapunov exponent λ as a function of σ^2

Fig. 2 The Lyapunov exponents λ as functions of the iteration by a conventional method and the proposed one ($N = 20$, $\lambda^{\text{obj}} = 0.2$)

によって得られた関係と比較した結果を Fig. 1 に示す。ただし、本研究では次節も含めて、各種定数を経験的に $M_2 = 5$, $\alpha = 1.2$, $\beta = 2$, $\gamma = 1.2$ と決定し、また、 $s_0 = 0.2$, $T = 100$ とした。これより、新手法が、 λ と σ^2 の関係の近似に成功していることがわかる。なお、 $N = 20, 200$ のネットワークに対して同様の近似精度が得られており、提案手法が統計的近似が有効な、ある程度以上のネットワークサイズに対してロバストであることが確認されている。

3.4 提案制御法

(21)式より、 $\partial\lambda/\partial\sigma^2$ は

$$\frac{\partial\lambda}{\partial\sigma^2} = \frac{1}{2} \left\{ \frac{1}{\sigma^2} + \frac{1}{E[\{X_i(t+1)\}^2]} \cdot \frac{\partial E[\{X_i(t+1)\}^2]}{\partial\sigma^2} \right\}, \quad (27)$$

であり、 $\partial E[\{X_i(t+1)\}^2]/\partial\sigma^2$ は、(26)式より容易に計算できる。したがって、(27)式を(14)式に代入することで、 $\partial x(t)/\partial w_{ij}$ の時系列データを必要とせず微係数 $\partial\lambda/\partial w_{ij}$ が近似的に求まる。

4. シミュレーション例

ネットワークの重み係数をランダムに初期化した完全結合型 RNN の最大リアプノフ指数を所望の値に制御する問題を考える。はじめに、3.1節で説明した従来法を用いて、 $N = 20$ の完全結合型 RNN の最大リアプノフ指数を $\lambda^{\text{obj}} = 0.2$ に制御するシミュレーション結果を Fig. 2 (破線) に示す。これより、最大リアプノフ指数が $\lambda \approx -0.05$ 付近で収束してしまい、ダイナミクスの制御に失敗していることがわかる。これは、 $\lambda^{\text{obj}} > 0$ の場合、カオスの初期値敏感性により微係数 $\partial x(t)/\partial w_{ij}$ の計算が発散してしまい δw_{ij} が正しく求まらないことが原因と考えられる。一方、3.4 節で提案した近似制御法を用いた場合の、同様のシミュレーションの結果を Fig. 2 (実線) に示す。これより、所望の最大リアプノフ指数が $\lambda^{\text{obj}} > 0$ の場合でもうまく制御できているのがわかる。なお、このシミュレーションで微係数 $\partial\lambda/\partial w_{ij}$ の計算に要した 1 回当たりの平均 CPU 時間は、前述の従来法の場合で 986 秒、後述の近似制御法の場合で 0.032 秒であった。簡単な比較ではあるが、構成ニューロン数が多い大規模なネットワークの最大リアプノフ指数を制御する場合、従来法を用いたの

Fig. 3 The Lyapunov exponent λ as a function of the iteration by the proposed approximation method ($N = 100$, $\lambda^{obj} = 1.0$)

では計算時間がかかりすぎるため実用的ではないことがわかる。一方、近似制御法を用いて、 $N = 100$ の完全結合型RNNの最大リアプノフ指数を $\lambda^{obj} = 1.0$ に設定するシミュレーション結果をFig. 3に示す。これより、 $N = 100$ くらいの大規模なネットワークを用いた場合でもうまく制御できていることがわかる。なお、この場合の平均CPU時間は0.094秒であった。

5. ま と め

本稿では、構成ニューロン数の多い完全結合型RNNの最大リアプノフ指数とニューロンの入力分散 σ^2 との関係の統計的解析を行った。その結果を基にして、最大リアプノフ指数を所望の値に制御するための近似的な定式化を行った。

シミュレーション結果より、この近似計算法が、構成ニューロン数の多い完全結合型RNNに対する高い計算コストを低減し、RNNがカオスダイナミクスをもつ場合でも精度の高い複雑さ制御が可能であることを示した。提案手法を用いてのダイナミクスの複雑さを考慮したRNNの学習についての考察は今後の課題である。

謝 辞

級数解析においてご助言頂きました東北大学医療技術短期大学部 洲之内 長一郎 教授に深く感謝いたします。

参 考 文 献

- 1) K. S. Narendra and K. Parthasarathy: Identification and Control of Dynamical Systems Using Neural Networks, IEEE Trans. Neural Networks, 1-1, 4/27 (1990)
- 2) J. J. Hopfield and D. W. Tank: "Neural" Computation of Decisions in Optimization Problems, Biological Cybernetics 52, 141/152 (1985)
- 3) 中野肇 他: ニューロコンピュータの基礎, コロナ社 (1990)
- 4) C. A. Skarda and W. J. Freeman: How brains make chaos in order to make sense of the world, Behavioral and Brain Sciences, 10, 161/195 (1987)
- 5) B. A. Pearlmutter: Learning State Space Trajectories in Recurrent Neural Networks, Neural Computation, 1-2, 263/269 (1989)
- 6) 長島寛之, 馬場良和: カオス入門, 培風館 (1992)
- 7) J. C. Principe and J. Kuo: Dynamic modelling chaotic time series with neural networks, In G. Tesauero et al.(Eds.), Neural Information Processing System 7, 311/318 (1995)

- 8) M. Jordan: Generic Constraints on Underspecified Target Trajectories, Proc. International Joint Conference on Neural Networks(IJCNN-89), 1, 217/225 (1989)
- 9) G. Deco and B. Schürmann: Dynamic modelling chaotic time series, Computational Learning Theory and Neural Learning Systems, 4 of Making learning Systems Practical, MIT Press, Cambridge, Mass., chapter 9, 137/153 (1997)
- 10) 本間, 酒井, 阿部, 竹田: 神経回路網ダイナミクスの複雑さ制御法, 計測自動制御学会論文集, 35-1, 138/143 (1999)
- 11) 酒井, 本間, 阿部: リカレントニューラルネットワークの複雑さの制御に関する一考察, 第41回自動制御連合講演会, 333/334 (1998)
- 12) K. Hirasawa, X. Wang, J. Murata, J. Hu and C. Jin: Universal learning network and its application to chaos control. Neural Networks, 13, 239/253 (2000)
- 13) A. Wolf, J. B. Swift, H. L. Swinney and J. A. Vastano: Determining Lyapunov Exponents from a time series, Physica, 16D, 285/315 (1989)
- 14) N. Honma, K. Kitagawa and K. Abe: Effect of Complexity on Learning Ability of Recurrent Neural Networks, Proc. of the 2nd AROB, 58/61 (1997)
- 15) 上田, 岡田, 芳谷: 確率と統計, 大日本図書 (1980)
- 16) 内田: 級数論上, 槇書店 (1969)

[著 者 紹 介]

酒 井 正 夫 (学生会員)

1997年, 東北大学工学部電子工学科卒業, 99年同大学大学院工学研究科博士前期課程修了。現在同大学大学院工学研究科博士後期課程に在学中。リカレントニューラルネットワークの学習に関する研究に従事。99年計測自動制御学会学術奨励賞, 2000年KACC Best Paper Award 受賞。

本 間 経 康 (正会員)

1995年, 東北大学大学院工学研究科博士後期課程修了。同年同大学電気通信研究所研究員, 96年同大学医療技術短期大学部講師, 98年同助教授, 2000年Saskatchewan大学(カナダ)客員教授。最適制御, 自律分散システム, ニューラルネットワーク, 人工生命などの研究に従事。IEEE, 日本放射線技術学会の会員(工学博士)。

阿 部 健 一 (正会員)

1969年, 東北大学大学院工学研究科博士課程修了。同年同大学工学部助手, 74年同大助教授, 84年豊橋技術科学大学教授, 93年東北大学教授。デジタル制御, 適応・学習制御などの研究に従事。IEEE, 電気学会, 情報処理学会などの会員(工学博士)。