複素リカレントニューラルネットワークの 大域的漸近安定性条件とその応用

非会員 吉田 光男* 正員 森 武宏*
正員 黒江 康明*

Global Asymptotic Stability Condition for Complex-valued Recurrent Neural Networks and Its Application

Yoshida Mitsuo*, Non-member, Takehiro Mori*, Member, Yasuaki Kuroe*, Member

In this paper, we propose a condition on global asymptotic stability for recurrent type complex-valued neural networks with a class of activation functions. Those networks have states, connection weights, and activation functions, which are all complex-valued. Such networks have been studied concerning their abilities of information processing, because of their attractive features which are not existent in real-valued counterparts. In particular, the activation function of a complex-valued neuron is an important factor that characterizes the dynamics of complex-valued neural networks. We propose a class of activation functions which ensure the global asymptotic stability, together with a condition on the connection weights. Furthermore, we apply the proposed dynamical networks and the stability condition to a convex programming problem with nonlinear constraints.

キーワード：複素リカレントニューラルネットワーク，大域的漸近安定性，凸計画問題

Keywords: complex-valued recurrent neural networks, global asymptotic stability, convex programming problem

1. はじめに

生体の神経系における情報処理の仕組みを模倣した人工ニューラルネットワークについては、様々なモデルが提案され、その振舞いの解析や応用についての研究が盛んである。そのなかで、ネットワークを複素数領域で表現し、複素数値情報を直接的に扱うようなモデルが信号処理や光工学の分野において提案されており、その能力について研究が行われている^{(1)~(4)}。この複素数値化されたニューラルネットワーク（複素ニューラルネットワーク）は、従来の実数領域におけるニューラルネットワーク（実数ニューラルネットワーク）を含む上、さらに実数ニューラルネットワークにはない特徴が存在するとの報告があり、ネットワークの情報処理能力に高い期待が持たれている⁽¹⁾。しかし、複素ニューラルネットワークでは、複素数値化されたニューロン（複素ニューロン）における信号の非線形変換処理を表現する複素関数（活性化関数）を如何に定義するかが、ネッ

トワークの振舞いを決定づけ、その能力を左右する大きな要素となっており、検討が待たれている。

本論文では、ネットワークにフィードバック構造を持つリカレントニューラルネットワークを複素数値化したモデル（複素リカレントニューラルネットワーク）を対象にする。実数領域におけるリカレントニューラルネットワークの応用では、ホップフィールドによる連想記憶の実現や最適化問題への適用が代表的である⁽⁵⁾⁽⁶⁾。このうち、最適化問題の解法へ適用する場合、ネットワークの大域的漸近安定性は、ネットワークが局所最適解を導かないための重要な性質として活発に研究がなされ理論的に明らかにされている^{(7)~(10)}。

一方、複素ニューラルネットワークでは、実数領域における連想記憶の実現問題を拡張して、複素領域でエネルギー関数を構成し、それを用いたネットワークの局所的な漸近安定性についての考察と、複素連想記憶の実現が試みられている⁽⁴⁾。しかし、最適化問題への応用の可能性やネットワークの大域的漸近安定性についての理論的な考察は未だ十分な研究がなされていないようである。そこで、本論文では、リアプノフ関数を導く複素ニューロンの活性化関数として一つの複素関数のクラスを定義し、そのネットワークの大域的漸近安定性条件を示す。また、その応用として、

* 京都工芸繊維大学 工芸学部 電子情報工学科
〒 606-8585 京都市左京区松ヶ崎
Department of Electronics and Information Science,
Kyoto Institute of Technology,
Matsugasaki Sakyo-ku Kyoto, 606-8585

得られた大域的漸近安定性条件を用いて解くことのできる最適化問題のクラスを示す。すなわち、ネットワークで凸計画問題が解けることを数値例と共に示す。

なお、以下では、 $\mathbb{C}^{n \times m}(\mathbb{R}^{n \times m})$ を複素 (実数) $n \times m$ 行列の全体、 $\text{Re}[u_i](\mathcal{I}[u_i])$ または $u_i^R(u_i^I)$ を複素数 u_i の実数部 (虚数部)、 $\bar{u}_i$ を複素数 u_i の共役複素数、添字 t は転置、 $*$ は共役転置とし、 $A > 0$ は行列 $A = A^*$ が正定値行列であることを表し、 $\lambda_{\min}(A)$ はエルミート行列 A の最小固有値、 $\|A\|_2$ はユークリッドノルムから導かれる行列 A のマトリクスノルムとする。また、虚数単位は j ($j^2 = -1$) である。

2. 複素ニューラルネットワークのモデル

本論文で対象とする複素ニューラルネットワークは、連続時間上で定義され、アナログ状態量をもつ複素ニューロンが相互に結合したネットワークであり、以下の複素数値変数を含む微分方程式で表される。

$$\begin{cases} \frac{du_i}{dt} = -d_i u_i + \sum_{k=1}^n w_{ik} v_k + b_i \\ v_i = f_i(u_i) \end{cases} \quad (i = 1, 2, \dots, n) \quad \dots (1)$$

ここで、 n はネットワークを構成するニューロンの個数、 $v_i \in \mathbb{C}$ 、 $u_i \in \mathbb{C}$ 、 $b_i \in \mathbb{C}$ はそれぞれ i 番目のニューロンの出力、膜電位、しきい値で、 $d_i \in \mathbb{R}$ は i 番目のニューロンの時定数を表す正のパラメーター、 $w_{ik} \in \mathbb{C}$ は k 番目から i 番目のニューロンへの結合重み、 $f_i: \mathbb{C} \rightarrow \mathbb{C}$ はニューロンの活性化関数とし、正の実数値である d_i 以外は全て複素数値をとる。

また、式 (1) のベクトル表現を以下のように定義する。

$$\begin{cases} \frac{du}{dt} = -Du + Wv + b \\ v = f(u) \end{cases} \quad \dots (2)$$

但し、 $W = \{w_{ik}\} \in \mathbb{C}^{n \times n}$ 、 $b = [b_1, b_2, \dots, b_n]^t \in \mathbb{C}^n$ 、 $u = [u_1, u_2, \dots, u_n]^t \in \mathbb{C}^n$ 、 $v = [v_1, v_2, \dots, v_n]^t \in \mathbb{C}^n$ 、 $f(u) = [f_1(u_1), f_2(u_2), \dots, f_n(u_n)]^t: \mathbb{C}^n \rightarrow \mathbb{C}^n$ 、 $D = \text{diag}\{d_1, d_2, \dots, d_n\} \in \mathbb{R}^n$ ($d_i > 0$) とおいている。

なお、本論文において、複素関数 $f_i(u_i): \mathbb{C} \rightarrow \mathbb{C}$ が偏微分可能であるとは、複素数を実数部と虚数部に分けて考え、

$$f_i(u_i) := f_i^R(u_i^R, u_i^I) + j f_i^I(u_i^R, u_i^I) \quad \dots (3)$$

と表し、実数値関数 $f_i^R, f_i^I: \mathbb{R}^2 \rightarrow \mathbb{R}$ が、それぞれ、実数値変数 $u_i^R, u_i^I \in \mathbb{R}$ について偏微分可能であることを意味し、その場合の偏導関数を、

$$\frac{\partial f_i^R}{\partial u_i^R}(\cdot), \quad \frac{\partial f_i^R}{\partial u_i^I}(\cdot), \quad \frac{\partial f_i^I}{\partial u_i^R}(\cdot), \quad \frac{\partial f_i^I}{\partial u_i^I}(\cdot) \quad \dots (4)$$

と書き表すこととする。

3. 複素ニューロンの活性化関数

複素ニューラルネットワーク (1) に対して、ネットワー

クに大域的漸近安定性をもたらす、最適化問題への適用を可能とする複素ニューロンの活性化関数を以下に定義する。

〔定義 1〕(クラス S) 複素関数 $f_i(u_i)$ は、実数値関数

$$\psi_i(r_i): \mathbb{R}_+ \rightarrow \mathbb{R}_+ \text{ を用いて、} \\ f_i(u_i) = \psi_i(|u_i|) \frac{u_i}{|u_i|} \quad \dots (5)$$

と書き表せ、この関数 $\psi_i(r_i)$ が以下の条件を満たす。

1. $|\psi_i(\cdot)|$ は有界である
2. $\psi_i(0) = 0$ を満たす
3. 連続な導関数 $\frac{d\psi_i}{dr_i}$ が $\forall r_i \geq 0$ で存在する
4. $\forall \xi \in \mathbb{R}, \exists l_i > 0, \text{ s.t. } 0 < \frac{d\psi_i}{dr_i}(\xi) \leq l_i < +\infty$

このような複素関数の集合をクラス S と定義する。

但し、 $\mathbb{R}_+ = \{r_i \mid r_i \in \mathbb{R}, r_i \geq 0\}$ としている。

定義 1 で、実数値関数 $\psi_i(\cdot)$ の条件 2 と、式 (5) における $u_i/|u_i|$ の部分が有界関数であることから、 $f_i(\cdot) \in S$ は、 $f_i(0) = 0$ であり、複素平面全体で定義される関数である。さらに、 $f_i(\cdot)$ は複素平面全体で偏微分可能である。なお、複素関数を式 (5) の形式に書き表せるとは、複素変数 u_i の偏角成分は保存し、その絶対値成分のみに非線形な変換が行われることを意味しており、クラス S に属する関数として、例えば以下の複素関数が挙げられる。

$$f_i(u_i) = \tanh(|u_i|) \frac{u_i}{|u_i|} \quad \dots (6)$$

$$f_i(u_i) = \frac{u_i}{1 + |u_i|} \quad \dots (7)$$

これらの複素関数は、これまでに複素ニューラルネットワークの活性化関数として用いられてきた関数である⁽²⁾⁻⁽⁴⁾。

また、クラス S に属する複素関数のネットワークの大域的漸近安定性を示すために必要となる性質を以下に示す。

〔性質 1〕ある固定された複素数を $u_i^e \in \mathbb{C}$ とし、関数 $h_i(z_i): \mathbb{C} \rightarrow \mathbb{C}$ を以下のように定義する。

$$h_i(z_i) := f_i(z_i + u_i^e) - f_i(u_i^e), \quad z_i \in \mathbb{C} \quad \dots (8)$$

この時、複素非線形関数 $f_i(\cdot)$ が、 $f_i(\cdot) \in S$ ならば、以下の等式を満足する。

$$\frac{\partial h_i^R}{\partial z_i^R} = \frac{\partial h_i^I}{\partial z_i^R} \quad \dots (9)$$

〔性質 2〕複素非線形関数 $f_i(\cdot) \in S$ は、任意の複素数 $a, b \in \mathbb{C}$ に対して、

$$\begin{aligned} & \text{Re} \left[\overline{\{f_i(a) - f_i(b)\}}(a - b) \right] \\ & \geq l_i^{-1} |f_i(a) - f_i(b)|^2 \quad \dots (10) \end{aligned}$$

を満足する。

性質 1 は、複素関数 $f_i(u_i)$ の偏導関数を計算することにより、

$$\frac{\partial f_i^R}{\partial u_i^R} = \frac{\partial f_i^I}{\partial u_i^R} = \left(\frac{d\psi_i}{dr_i}(|u_i|) - \frac{\psi_i(|u_i|)}{|u_i|} \right) \frac{u_i^R u_i^I}{|u_i|^2} \quad \dots (11)$$

が成り立つことで導くことができる。また、性質 2 の導出は付録に記す。

4. ネットワークの大域的漸近安定性条件

複素ニューラルネットワークの活性化関数として、クラス $\mathcal{S}$ に属する関数を用いた場合、以下の定理のように、結合重みに適当な条件を課すことで、ネットワークの大域的漸近安定性を保証することができる。

〔定理 1〕 複素ニューロンの活性化関数が $f_i(u_i) \in \mathcal{S}$ ($i = 1, 2, \dots, n$) である複素ニューラルネットワーク (1) において、 $L = \text{diag}\{l_1^{-1}, l_2^{-1}, \dots, l_n^{-1}\} \in \mathbb{R}^{n \times n}$ とおいた時、結合重み W に対して、

$$P(DL - W) + (DL - W)^*P > 0 \dots\dots\dots (12)$$

を満たす正定値対角行列 $P \in \mathbb{R}^{n \times n}$ が存在するならば、複素ニューラルネットワークの平衡点は存在し、かつ大域的漸近安定である。

定理 1 の証明 平衡点の存在は、ニューロンの非線形関数が有界関数であることから実数ニューラルネットワークの場合⁽⁹⁾と同様に示すことができる。すなわち、平衡点の存在は、写像 $\chi(u) := D^{-1}(Wf(u) + b)$ の不動点の存在問題であり、Brower の不動点定理「連続写像の像が有界ならば、写像は不動点を持つ」⁽¹¹⁾より、 $\chi(\cdot)$ は有界なので不動点を持ち平衡点の存在が保証される。

平衡点が大域的漸近安定であることは、存在する平衡点の一つを $u^e \in \mathbb{C}^n$ とおき、ネットワーク (2) を $z = u - u^e$ により、以下のように変数変換する。

$$\begin{cases} \frac{dz}{dt} = -Dz + Wh(z) \\ h(z) = f(z + u^e) - f(u^e) \end{cases} \dots\dots\dots (13)$$

このシステムに対して、原点が大域的漸近安定な平衡点であることをリアプノフ関数を用いて示す。構成するリアプノフ関数は次のような関数である。

$$V(z) = \frac{1}{2} z^* D^{-1} z + \frac{k}{\epsilon} \sum_{i=1}^n p_i G_i(z_i^R, z_i^I) \dots\dots\dots (14)$$

但し、 p_i ($i = 1, 2, \dots, n$) は定理の条件 (12) を満たす正定値対角行列 P の対角要素とし、定数 ϵ, k は次式で与えられるものである。

$$\begin{cases} \epsilon \in (0, 1) \\ k = \|D^{-1}W\|_2^2 / 4\lambda_{\min}(Q) > 0 \\ 2Q = P(DL - W) + (DL - W)^*P \end{cases} \dots\dots\dots (15)$$

さらに、 $G_i(\cdot, \cdot)$ は、以下のように定義されるものである。

$$G_i(z_i^R, z_i^I) = \int_0^{z_i^R} h_i^R(\rho, 0) d\rho + \int_0^{z_i^I} h_i^I(z_i^R, \rho) d\rho \dots\dots\dots (16)$$

ここで、式 (16) の関数は、非線形関数 $f_i(\cdot) \in \mathcal{S}$ の性質 1 より、以下のように変形でき、

$$G_i(z_i^R, z_i^I) = \int_0^1 \text{Re}\{\bar{z}_i h_i(\eta z_i)\} d\eta \geq 0 \dots\dots\dots (17)$$

これと、性質 2 より $G_i(\cdot, \cdot)$ は非負であるので、リアプノ

フ関数 (14) は、 $V(z) \geq (1/2) \cdot z^* D^{-1} z$ となる。

次に、 $V(z(t))$ を $V(z^R(t), z^I(t)) : \mathbb{R} \rightarrow \mathbb{R}^{2n} \rightarrow \mathbb{R}$ とみなし、ネットワーク (13) の解に沿って時間微分する。関数 $h_i(z_i)$ ($i = 1, 2, \dots, n$) が性質 1 を満たすことから、

$$\begin{aligned} \frac{dG_i}{dt} \Big|_{(13)} &= \frac{\partial G_i}{\partial z_i^R} \frac{dz_i^R}{dt} + \frac{\partial G_i}{\partial z_i^I} \frac{dz_i^I}{dt} \\ &= h_i^R(z_i^R, z_i^I) \frac{dz_i^R}{dt} + h_i^I(z_i^R, z_i^I) \frac{dz_i^I}{dt} \\ &= \text{Re}\left\{ \bar{h}_i(z_i) \frac{dz_i}{dt} \right\} \dots\dots\dots (18) \end{aligned}$$

であるので、次の関係が成り立つ。

$$\begin{aligned} \frac{dV(z)}{dt} \Big|_{(13)} &= \sum_{i=1}^n \frac{\partial V}{\partial z_i^R} \frac{dz_i^R}{dt} + \frac{\partial V}{\partial z_i^I} \frac{dz_i^I}{dt} \\ &= \text{Re}\left\{ \left[D^{-1}z + \frac{k}{\epsilon} Ph(z) \right]^* \frac{dz}{dt} \right\} \\ &= -\|z\|_2^2 + \text{Re}\left\{ z^* D^{-1} Wh(z) \right\} \\ &\quad - \frac{k}{\epsilon} \text{Re}\left\{ z^* PDh(z) - h(z)^* PWh(z) \right\} \\ &\leq -\|z\|_2^2 + \text{Re}\left\{ z^* D^{-1} Wh(z) \right\} \\ &\quad - \frac{k}{\epsilon} \text{Re}\left\{ h(z)^* PDLh(z) - h(z)^* PWh(z) \right\} \\ &= -\|z\|_2^2 + \text{Re}\left\{ z^* D^{-1} Wh(z) \right\} - \frac{k}{\epsilon} h(z)^* Qh(z) \dots\dots\dots (19) \end{aligned}$$

ここで、上式の不等式には、性質 2 より導かれる次の関係式を用いている。

$$\text{Re}\left\{ z^* PDh(z) \right\} \geq h(z)^* PDLh(z) \dots\dots\dots (20)$$

さらに、 $\epsilon \in (0, 1)$ より、

$$\text{Re}\left\{ z^* D^{-1} Wh(z) \right\} \leq \epsilon \|z\|_2^2 + \frac{\|D^{-1}Wh(z)\|_2^2}{4\epsilon} \dots\dots\dots (21)$$

の関係が成り立ち、これを用いると、結局 $V(z)$ の時間微分は以下のように評価される。

$$\begin{aligned} \frac{dV(z)}{dt} \Big|_{(13)} &\leq -(1 - \epsilon) \|z\|_2^2 \\ &\quad + \left(\frac{\|D^{-1}W\|_2^2}{4\epsilon} - \frac{k\lambda_{\min}(Q)}{\epsilon} \right) \|h(z)\|_2^2 \\ &= -(1 - \epsilon) \|z\|_2^2 \dots\dots\dots (22) \end{aligned}$$

よって、 $dV(z)/dt|_{(13)} \leq 0$ であり、等号が成り立つのは原点のみなので、 $z = 0$ 、すなわち平衡点は漸近安定である。さらに、リアプノフ関数 (14) は、行列 D が正定値対角行列、関数 $G_i(\cdot, \cdot)$ ($i = 1, 2, \dots, n$) が正定値関数なので、 $\|z\| \rightarrow \infty$ の時、 $\|V(z)\| \rightarrow \infty$ となり、平衡点の漸近安定性は大域的である。□

5. 凸計画問題への適用

リカレント型の実数ニューラルネットワークでは、その応用として、連想記憶と並び、最適化問題への適用は代表的である⁽⁵⁾⁻⁽⁷⁾。ここでは、本論文で対象としている複素ニューラルネットワークの最適化問題への適用例を示す。対象とする問題は、 $M \in \mathbb{C}^{n \times n}$ を与えられたエルミートな正定値行列として、

$$\begin{aligned} \min_v \quad & J(v) = \frac{1}{2} v^* M v - \operatorname{Re}(q^* v) \quad \dots\dots\dots (23) \\ \text{s.t.} \quad & |v_i| \leq K_i \quad (i = 1, 2, \dots, n) \\ & v = [v_1, v_2, \dots, v_n]^t \in \mathbb{C}^n \\ & q = [q_1, q_2, \dots, q_n]^t \in \mathbb{C}^n \end{aligned}$$

で定式化される問題である。この問題の制約条件は、複素変数 v_i の絶対値に関して上限を定める制約であり、凸なる領域内で凸なるエルミート形式の目的関数を最小化する問題である。また、複素数を実数部と虚数部で構成される二つの実数の並びとみなし、問題 (23) を実数領域における問題に変換すれば、その制約条件は非線形不等式 ($\sqrt{v_i^2 + v_i'^2} \leq 1$) であるので、非線形な制約条件が課せられた凸計画問題とみなせる。この種の問題は、通常の二次計画問題と比べて、標準的な解法は十分整備されておらず、なお改良の余地のある問題である。

この問題に対して、以下の複素ニューラルネットワークを考えると、

$$\begin{cases} \frac{du}{dt} = -Du + (-M)f(u) + q \\ f_i(u_i) = \psi_i(|u_i|) \frac{u_i}{|u_i|} \in \mathcal{S} \\ |\psi_i(\cdot)| \leq K_i \quad (i = 1, 2, \dots, n) \end{cases} \quad \dots\dots\dots (24)$$

実数ニューラルネットワークと同様に、ネットワーク (24) のエネルギー関数を定義することによって、平衡点として問題 (23) の近似解を見通し良く与えることができる。その概略は以下のとおりである。

複素ニューラルネットワーク (24) に対して、次のエネルギー関数 $E(v)$ を定義する。

$$\begin{aligned} E(v) = & \frac{1}{2} v^* M v - \operatorname{Re}\{q^* v\} \\ & + \sum_{i=1}^n d_i \int_0^{|v_i|} \psi_i^{-1}(\rho) d\rho \quad \dots\dots\dots (25) \end{aligned}$$

ここで、 $\psi_i^{-1}(\cdot)$ は実数値関数 $\psi_i(\cdot)$ の逆関数であり、 $\psi_i(0) = 0$ であることから、単調増加関数であることから、式 (25) の積分項は非負である。さらに、行列 M は正定値エルミート行列であることから、エネルギー関数 $E(v)$ は下に有界な関数である。

また、結合重み行列 $W = (-M)$ がエルミート行列で非線形関数が $f_i(u_i) \in \mathcal{S}$ の時、ネットワーク (24) の解に沿った時間微分は、 $r_i = |u_i|$ 、 $\theta_i = \arg(u_i)$ とおくと、

$$\begin{aligned} \frac{dE}{dt} \Big|_{(24)} = & -d_i \sum_{i=1}^n \left[\frac{d\psi_i}{dr_i} \left(\frac{dr_i}{dt} \right)^2 + r_i \psi(r_i) \left(\frac{d\theta_i}{dt} \right)^2 \right] \leq 0 \\ & \dots\dots\dots (26) \end{aligned}$$

と計算され、等号が成立するのは $dr_i/dt = 0$ 、 $d\theta_i/dt = 0$ ($i = 1, 2, \dots, n$) の時のみである。よって、ネットワーク (24) は、エネルギー関数 (25) を減少させるように動作する。

ここで、行列 M はエルミート正定値行列であるので、複素ニューラルネットワーク (24) の結合重み行列 $-M$ は定理 1 の条件を満たす。従って、与えられた任意の初期条件に対するネットワークの全ての解は大域的漸近安定な平衡点へ収束し、その平衡点はエネルギー関数の最小点である。

また、エネルギー関数 (25) は、問題 (23) を非線形関数 $\int_0^{|v_i|} \psi_i^{-1}(\rho) d\rho$ を用いて、制約のない最小化問題の目的関数へ変換したものであり、第 3 項はペナルティ項、 $d_i > 0$ ($i = 1, 2, \dots, n$) はペナルティパラメーターに相当し、パラメーター $d_i > 0$ ($i = 1, 2, \dots, n$) を調整することによってエネルギー関数の最小点を問題 (23) の最適解へ十分近づけることができる。よって、複素ニューラルネットワーク (24) により、凸計画問題 (23) の近似解を求めることが可能である。ちなみに、このエネルギー関数 (25) は文献 (2) で定義されている関数と同じ形であり。式 (26) も文献 (2) で計算されている。

一般に、複素変数を含む最小化問題は、複素数を二次元の実数ベクトルとみなして、実数上の最小化問題として扱える。したがって、最小化問題を実数上の問題とみなして、実数ニューラルネットワークを適用することも可能であるとも考えられる。しかし、本論文で扱った最小化問題は、問題の許容領域が複素平面上の円領域であり、実数空間上の超立方体領域内で最小化する従来の実数ニューラルネットワーク⁽⁵⁾⁻⁽⁷⁾は直接的に適用できない。一方で、複素ニューラルネットワークでは、実数の場合と同じ論法により、そのような問題が適用可能となる。つまり、ネットワークの複素数値化により、ニューラルネットワークが簡単な形で扱える最適化問題のクラスを広げることができる。

6. 数値例

凸計画問題 (23) を実際に複素ニューラルネットワークを用いて解いた数値例を示す。設定値は以下の通りである。

$$\begin{aligned} M = & \begin{pmatrix} 3.0 & 1.0 - j1.0 \\ 1.0 + j1.0 & 4.0 \end{pmatrix}, \quad q = \begin{pmatrix} 15 + j5.0 \\ 20 + j20 \end{pmatrix} \\ K_1 = & 1.0, \quad K_2 = 1.0 \end{aligned}$$

なお、最適解は、 $(v_1^{opt}, v_2^{opt}) = (0.94 + j0.35, 0.72 + j0.69)$ である。この問題に対して複素ニューラルネットワーク (24) を適用する。適用するネットワークのパラメーターは、 $D = \operatorname{diag}\{0.01, 0.01\}$ 、 $W = (-M)$ 、 $b = q$ であり、ニューロンの非線形関数は式 (7) の関数を用いた。

この時、複素ニューラルネットワークの重み行列 W はエルミート行列であり、その固有値は、 -2.0 と -5.0 であ

図1 複素平面上の最適点へ収束するネットワークの出力 $v_i \in \mathbb{C}$ の軌道

Fig. 1. Trajectories of output of neural network, v_1 , converging on optimum point in complex plane.

るので、大域的漸近安定条件を満たしている。

このニューラルネットワークを実際に動作させた結果、 $(v_1, v_2) = (0.938 + j0.347, 0.720 + j0.694)$ に収束し、問題の近似解が得られた。図1は、幾つかの制約条件を満たす初期点から複素ニューラルネットワークを動作させた時のネットワークの出力 $v_1 \in \mathbb{C}$ の軌道を示している。

7. おわりに

本論文では、複素リカレントニューラルネットワークに対して、その定性的性質を規定するのに重要な役割を果たすニューロンの活性化関数に関して一つのクラスを定め、活性化関数がそのクラスに属するような複素ニューラルネットワークについて、平衡点が大域的漸近安定であるための十分条件を提案した。定めた複素ニューロンの活性化関数のクラスは、幾つかのこれまでに提案されている活性化関数を含んでおり、それらの大域的漸近安定性条件が得られたことになっている。また、本論文の結果は、複素数の虚数部を0とおいた場合、実数ニューラルネットワークにおける結果⁽⁹⁾⁽¹⁰⁾の安定性条件に対応させることが可能であり、それらの結果の拡張であるといえる。

さらに、ニューロンの活性化関数が本論文で定義したクラスに属するような複素ニューラルネットワークに対して、その大域的漸近安定性条件を用いて解くことのできる最適化問題のクラスを示した。示したクラスの凸計画問題は、非線形な制約条件を持ち、従来の実数ニューラルネットワークでは直接的に扱うことのできない問題のクラスである。

また、複素リカレントニューラルネットワーク自身についても、複素ニューロンの活性化関数の定義に関する問題をはじめ、実数ニューラルネットワークと異なる性質の存在

など未だ明らかになっていない部分が多い。本論文における大域的漸近安定性条件や最適化問題への適用例は、その端緒に過ぎない。例えば、本論文のクラス S の非線形関数は、その絶対値に関して、実数値のシグモイド関数の性質を受け継ぎ、実数ニューラルネットワークの成果を継承しようとしたものである。今後、様々なクラスの活性化関数について、複素ニューラルネットワークの性質や能力を検証していきたい。

(平成15年11月19日受付、平成16年5月6日再受付)

文 献

- (1) T. Nitta: "Complex-valued neural networks", *Journal of IEICE Japan*, Vol.83, No.8, pp.612-615 (2000-8) (in Japanese)
新田 徹:「複素ニューラルネットワーク」, 信学論誌, **83**, 8, pp.612-615 (2000-8)
- (2) M. Takeda and T. Kishigami: "Complex neural fields with a Hopfield-like energy function and an analogy to optical fields generated in phase-conjugate resonators", *J. Opt. Soc. Am. A.*, Vol.9, No.12, pp.2182-2191 (1992)
- (3) K. Onishi and A. Hirose: "Dynamics of recurrent complex-valued neural networks", *Tech. Rep. IEICE Neurocomputing*, NC96-123, pp.79-86 (1997) (in Japanese)
大西啓史・廣瀬 明:「複素リカレントニューラルネットワークのダイナミクス」, 信学技報, NC96-123, pp.79-86 (1997)
- (4) N. Hashimoto, Y. Kuroe, and T. Mori: "On energy function for complex-valued neural networks", *Trans. of ISCIE*, Vol.15, No.10, pp.559-565 (2002) (in Japanese)
橋本直樹・黒江康明・森 武宏:「複素ニューラルネットワークのエネルギー関数について」, システム制御情報学会論誌, **15**, 10, pp.559-565 (2002-10)
- (5) J.J. Hopfield: "Neurons with graded response have collective computation properties like those of two-state neurons", *Proc. Natl Acad. Sci. USA*, Vol.81, pp.3088-3092, May (1984)
- (6) D.W. Tank and J.J. Hopfield: "Simple 'neural' optimization networks: An A/D converter, signal decision circuit, and a linear programming network", *IEEE Trans. Circuits Syst. I*, Vol.33, pp.533-541 (1986)
- (7) M. Forti, S. Manetti, and M. Marini: "A condition for global convergence of a class of symmetric neural circuits", *IEEE Trans. Circuits Syst. I*, Vol.39, pp.480-483 (1992)
- (8) M. Forti, S. Manetti, and M. Marini: "Necessary and sufficient condition for absolute stability of neural networks", *IEEE Trans. Circuits Syst. I*, Vol.41, pp.491-494 (1994)
- (9) M. Forti and A. Tesi: "New conditions for global stability of neural networks with application to linear and quadratic programming problems", *IEEE Trans. Circuits Syst. I*, Vol.42, pp.354-366 (1995)
- (10) X.B. Liang and J. Si: "Global exponential stability of neural networks with globally Lipschitz continuous activations and its application to linear variational inequality problem", *IEEE Trans. Neural Networks*, Vol.12, pp.349-359 (2001)
- (11) 中岡 稔: 不動点定理とその周辺, 岩波書店, pp.7-14 (1977)

付 録

1. 性質2の導出

任意の複素数 $a, b \in \mathbb{C}$ が等しい場合は自明であり、 $a \neq b$ の場合を考える。複素数 a, b を $a = r_1 e^{j\theta_1}$, $b = r_2 e^{j\theta_2}$ と極形式で表し、(i) $r_1 \neq r_2$ の場合と、(ii) $r_1 = r_2$ かつ $\theta_1 \neq \theta_2$ の場合のそれぞれについて性質2の不等式を導く。

i. $r_1 \neq r_2$ の場合 実数値関数 $\psi_i(\cdot)$ は以下の不等式を満たす。

$$0 < \frac{\psi(r_1) - \psi(r_2)}{r_1 - r_2} \leq l_i < +\infty \dots\dots\dots (\text{付 1})$$

逆数をとる、各辺に $[\psi(r_1) - \psi(r_2)]^2$ を乗ずると、

$$l_i^{-1} [\psi(r_1) - \psi(r_2)]^2 \leq [\psi(r_1) - \psi(r_2)](r_1 - r_2) \dots\dots (\text{付 2})$$

を得る。また、 $\psi(0) = 0$ より、 $l_i^{-1}\psi(r_k) \leq r_k$, ($k = 1, 2$) を満たす。これらをふまえて式 (10) の左辺を以下のように計算することで性質 2 の不等式が導かれる。

$$\begin{aligned} \operatorname{Re} \left[\overline{\{f_i(a) - f_i(b)\}}(a - b) \right] &= r_1\psi(r_1) + r_2\psi(r_2) \\ &\quad - \{r_1\psi(r_2) + r_2\psi(r_1)\} \cos(\theta_1 - \theta_2) \\ &= \{\psi(r_1) - \psi(r_2)\}(r_1 - r_2) \\ &\quad + \{r_1\psi(r_2) + r_2\psi(r_1)\} \\ &\quad \times \{1 - \cos(\theta_1 - \theta_2)\} \\ &\dots\dots\dots (\text{付 3}) \end{aligned}$$

(付 3) 式の第 1 項に不等式 (付 2) を用いると、

$$\begin{aligned} (\text{上式}) &\geq l_i^{-1} [\psi(r_1) - \psi(r_2)]^2 \\ &\quad + \{r_1\psi(r_2) + r_2\psi(r_1)\} \\ &\quad \times \{1 - \cos(\theta_1 - \theta_2)\} \\ &= l_i^{-1} |\psi(r_1)e^{j\theta_1} - \psi(r_2)e^{j\theta_2}|^2 \\ &\quad + [\{r_1 - l_i^{-1}\psi(r_1)\}\psi(r_2) \\ &\quad + \{r_2 - l_i^{-1}\psi(r_2)\}\psi(r_1)] \\ &\quad \times \{1 - \cos(\theta_1 - \theta_2)\} \\ &\geq l_i^{-1} |f_i(a) - f_i(b)|^2 \dots\dots\dots (\text{付 4}) \end{aligned}$$

ここで、最後の不等式は $1 - \cos(\theta_1 - \theta_2) \geq 0$ と $l_i^{-1}\psi(r_k) \leq r_k$ ($k = 1, 2$) を用いている。

ii. $r_1 = r_2$ かつ $\theta_1 \neq \theta_2$ の場合 式 (10) の左辺を計算し、 $l_i^{-1}\psi(r_1) \leq r_1$ であることを用いると、

$$\begin{aligned} \operatorname{Re} \left[\overline{\{f_i(a) - f_i(b)\}}(a - b) \right] &= r_1\psi(r_1)|e^{j\theta_1} - e^{j\theta_2}|^2 \\ &\geq l_i^{-1}\psi(r_1)^2|e^{j\theta_1} - e^{j\theta_2}|^2 \\ &= l_i^{-1}|\psi(r_1)e^{j\theta_1} - \psi(r_1)e^{j\theta_2}|^2 \dots (\text{付 5}) \end{aligned}$$

さらに $r_1 = r_2$ より、

$$\begin{aligned} (\text{上式}) &= l_i^{-1}|\psi(r_1)e^{j\theta_1} - \psi(r_2)e^{j\theta_2}|^2 \\ &= l_i^{-1}|f_i(a) - f_i(b)|^2 \dots\dots\dots (\text{付 6}) \end{aligned}$$

となる。よって、

$$\operatorname{Re} \left[\overline{\{f_i(a) - f_i(b)\}}(a - b) \right] \geq l_i^{-1}|f_i(a) - f_i(b)|^2$$

が導かれた。□

吉 田 光 男 (非会員) 1978 年生。2001 年京都工芸繊維大学工芸学部電子情報工学科卒業、2003 年同大学大学院工芸科学研究科電子情報工学専攻博士前期課程修了。同年同博士後期課程進学、現在に至る。

森 武 宏 (正員) 1945 年生。1968 年京都大学工学部電子工学科卒業。1974 年同大学大学院工学研究科博士課程電気工学専攻修了。同年同大学工学部オートメーション研究施設助手、同応用システム科学教室助教を経て、現在、京都工芸繊維大学工芸学部電子情報工学科教授。この間、主として制御系の安定解析に関する研究に従事。工学博士。IEEE、計測自動制御学会、電子情報通信学会各会員。

黒 江 康 明 (正員) 1950 年生。1982 年神戸大学大学院自然科学研究科博士課程修了。同年神戸大学工学部電気工学科助手、1991 年京都工芸繊維大学工芸学部電子情報工学科助教授、2003 年同大学情報科学センター教授。システム制御理論とその応用、ニューロコンピューティングなどの研究に従事。学術博士。IEEE、計測自動制御学会、日本ロボット学会、電子情報通信学会各会員。