

Title	3ニューロンモデルにおける過渡現象への構造の影響 (生物数学の理論とその応用)
Author(s)	福井, 利彦; 芦澤, 恵太; 宮崎, 倫子
Citation	数理解析研究所講究録 (2005), 1432: 105-110
Issue Date	2005-05
URL	http://hdl.handle.net/2433/47391
Right	
Type	Departmental Bulletin Paper
Textversion	publisher

3 ニューロンモデルにおける過渡現象への構造の影響

静岡大学・大学院・理工学研究科 福井利彦 (Toshihiko Fukui), 芦澤恵太 (Keita Ashizawa)
Graduate School of Science and Technology, Shizuoka University

静岡大学・工学部 宮崎倫子 (Rinko Miyazaki)
Faculty of Engineering, Shizuoka University

1 はじめに

生物内のニューラルネットワークにおいても、人工ニューラルネットワークにおいても各ニューロンの情報伝達及び情報処理において時間遅れを考えることは自然なことである [1-4]. そして、多くの力学系において時間遅れは不安定性をもたらす要因として知られているように、ニューラルネットワークモデルにおいてもネットワークの収束性の喪失の原因となることが知られている. 例えば, Marcus and Westervelt [3] は、(時間遅れがなければ収束することによく知られている) 連続値ホップフィールドモデルにおいて時間遅れを導入すると平衡点への収束性は崩れ解が振動することを示している. 一方、各ニューロンの結合が全て興奮性の結合である場合には、単調性の理論 [5] が適用でき、たとえ時間遅れが存在しても平衡点への収束性が保たれることは分っている. しかし、Pakdaman et. al. [4] は、興奮性ニューロンが環状で一方向に接続されたモデル (つまり、平衡点への収束性が保たれるような場合) において、数値計算で観察される非常に長い過渡的振動が時間遅れによってもたらされることを指摘し解析を行っている. 彼らの結果は、シミュレーションの世界と理論的解析によって保証されている事実とのギャップを指摘した興味深い結果である. 我々は、彼らが指摘した長い過渡的な振動がネットワーク構造を変えた場合にも起こるのかどうかということに興味を持ち、簡単のため3つのニューロンからなるネットワークモデルを考えその解析を行った. 本稿では次の時間遅れを持つ微分方程式

$$\frac{dx}{dt}(t) = \frac{1}{\epsilon} [-x(t) + WBF(x(t-\tau))] \quad (\text{E})$$

を考える. ここで $x(t) = \text{col}(x_1(t), x_2(t), x_3(t))$ とし $x_i(t)$ ($i = 1, 2, 3$) は i 番目のニューロンの内部状態を表し、 $\epsilon > 0$ は時定数を表す. また各ニューロンの出力関数は $F(x) = \text{col}(\tanh x_1, \tanh x_2, \tanh x_3)$ で与えられ、結合荷重 $W > 0$ および 3×3 の接続行列 $B = (b_{ij})$ によって各ニューロンの結合の状態を表すものとする. B の非対角成分が非負 (興奮性ニューロンのみからなるネットワーク) でかつ行列 B が既約である時には、先にも述べたように単調性理論が適用でき、(E) のほとんど全ての解が平衡点に収束することが分っている. なお、接続行列 B を

$$B = B_1 := \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

で与えると、Pakdaman らの結果が適用でき、時定数 ϵ が小さいときには長く続く過渡的な振動が数値的に観測される. 彼らはこの過渡的振動を $\epsilon = 0$ とすることによって得られる差分方程式を用いて説明している. 我々は、 $B = B_1$ 以外に、新たに次の2つのネットワーク構造を考える. 図1は $B = B_1, B_2, B_3$ のそ

れぞれの場合についてのネットワークの概念図である。

$$B = B_2 := \frac{1}{2} \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \quad B = B_3 := \frac{1}{3} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

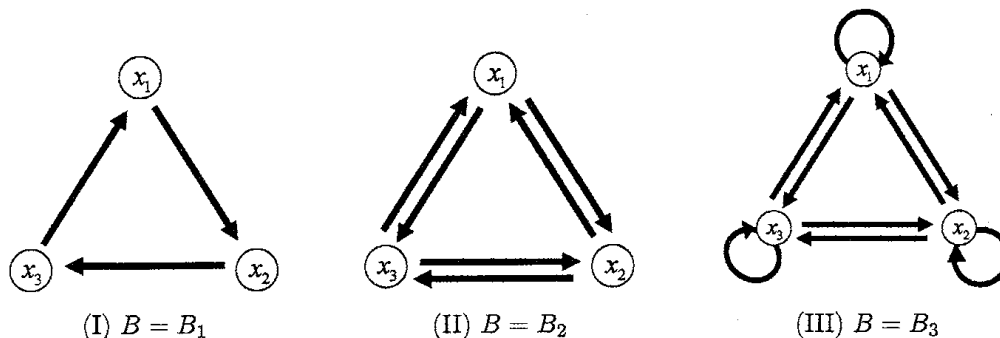


図 1: 3 種類のネットワークの概念図

2 漸近挙動

まず、通常の平衡点まわりでの局所的な安定性解析を行うことにより次の定理が得られる。

定理 1 (局所安定性).

(E) において、接続行列 B が B_1 , B_2 もしくは B_3 で与えられたとき、 $W < 1$ ならば (E) の平衡点は原点のみであり、原点は漸近安定である。 $W > 1$ ならば (E) は原点に対称な 2 点 $\pm(a, a, a)$ を加えた 3 つの平衡点を持ち、原点は不安定、他の 2 点は漸近安定である。ここで、 a は $a = W \tanh a$ の正の解である。

次に、方程式 (E) が共生系であることに注意し、単調性理論 [5] を用いることにより次の定理が得られる。

定理 2 (大域安定性).

(E) において、接続行列 B が B_1 , B_2 もしくは B_3 で与えられたとき、 $W < 1$ ならば、原点は大域的漸近安定である。 $W > 1$ ならば、ほとんどすべての解は (a, a, a) または $-(a, a, a)$ のいずれかに収束する。

以下では $W > 1$ の場合を考え、安定平衡点 (a, a, a) に収束する解がどのような挙動を呈するのかをより詳細に解析してみよう。なお、 $-(a, a, a)$ についても全く同様の議論が成り立つ。 (a, a, a) まわりでの線形化方程式は、

$$\frac{dy}{dt}(t) = \frac{1}{\epsilon} [-y(t) + g'(a)By(t - \tau)]. \quad (1)$$

ここで、 $g(x) = W \tanh x$ とおいた。したがって、 $g'(a) = \frac{4W}{(e^a + e^{-a})^2} > 0$ であることに注意しよう。 $B = B_1, B_2, B_3$ の各場合において、(1) に対する特性方程式は以下の通り求められる。

(I) $B = B_1$ のとき,

$$\begin{aligned} P_1(\lambda) &:= \left(\lambda + \frac{1}{\epsilon}\right)^3 + \left(-\frac{1}{\epsilon}e^{-\lambda\tau}g'(a)\right)^3 \\ &= \left(\lambda + \frac{1}{\epsilon} - \frac{g'(a)e^{-\lambda\tau}}{\epsilon}\right) \left(\lambda + \frac{1}{\epsilon} - \frac{g'(a)e^{-\lambda\tau}}{\epsilon}e^{\frac{2\pi i}{3}}\right) \left(\lambda + \frac{1}{\epsilon} - \frac{g'(a)e^{-\lambda\tau}}{\epsilon}e^{-\frac{2\pi i}{3}}\right) = 0. \end{aligned}$$

(II) $B = B_2$ のとき,

$$\begin{aligned} P_2(\lambda) &:= \left(\lambda + \frac{1}{\epsilon}\right)^3 - 3\left(\lambda + \frac{1}{\epsilon}\right) \left(\frac{1}{2\epsilon}e^{-\lambda\tau}g'(a)\right)^2 - 2\left(\frac{1}{2\epsilon}e^{-\lambda\tau}g'(a)\right)^3 \\ &= \left(\lambda + \frac{1}{\epsilon} - \frac{g'(a)e^{-\lambda\tau}}{\epsilon}\right) \left(\lambda + \frac{1}{\epsilon} + \frac{g'(a)e^{-\lambda\tau}}{2\epsilon}\right)^2 = 0. \end{aligned}$$

(III) $B = B_3$ のとき,

$$\begin{aligned} P_3(\lambda) &:= \left(\lambda + \frac{1}{\epsilon}\right)^3 + 3\left(\lambda + \frac{1}{\epsilon}\right)^2 \left(-\frac{1}{3\epsilon}e^{-\lambda\tau}g'(a)\right) \\ &= \left(\lambda + \frac{1}{\epsilon} - \frac{g'(a)e^{-\lambda\tau}}{\epsilon}\right) \left(\lambda + \frac{1}{\epsilon}\right) = 0. \end{aligned}$$

いずれの場合においても, 共通因数,

$$q_0(\lambda) := \lambda + \frac{1}{\epsilon} - \frac{g'(a)e^{-\lambda\tau}}{\epsilon}$$

が含まれていることに注意しよう. 共通因数以外の因数を次のようにおくこととする.

$$q_1(\lambda) := q_1^+(\lambda)q_1^-(\lambda), \quad q_1^\pm(\lambda) := \lambda + \frac{1}{\epsilon} - \frac{g'(a)e^{-\lambda\tau}}{\epsilon}e^{\pm\frac{2\pi i}{3}} \quad (\text{複号同順}),$$

$$q_2(\lambda) := \lambda + \frac{1}{\epsilon} + \frac{g'(a)e^{-\lambda\tau}}{2\epsilon},$$

$$q_3(\lambda) := \lambda + \frac{1}{\epsilon}.$$

補題 3.

q_0 から得られる固有値のうち実部最大のものを λ_0 とすると, λ_0 は負の実数で, q_0 以外の因数 q_1, q_2, q_3 から得られる任意の固有値 λ に対して $\operatorname{Re} \lambda < \lambda_0$ となる.

補題 3 で与えられた固有値 λ_0 は解の平衡点への収束を支配するので, λ_0 のことを以後支配的固有値と呼ぶこととする.

補題 4.

q_0 から得られる任意の固有値 λ (支配的固有値 λ_0 も含む) に対して λ に属する固有関数は $e^{\lambda s}v_0$, ($-\tau \leq s \leq 0$) で与えられる. ただし $v_0 = \operatorname{col}(1, 1, 1)$.

方程式(E)の解 $x(t)$ の挙動を3次元座標空間 $x_1x_2x_3$ にグラフ化したものを考えよう. そこで, $x_1x_2x_3$ 座標空間において, 平衡点 (a, a, a) を通り方向ベクトル v_0 をもつ直線を直線 l と呼ぶことにする. 補題 3, 4より次の定理が得られる.

定理 5 (漸近的挙動).

(E)において重み接続行列 B が B_1, B_2 もしくは B_3 で与えられたとき, 平衡点 (a, a, a) に収束する解は $t \rightarrow \infty$ のとき, 直線 l に沿って $e^{\lambda_0 t}$ の割合で指数的に収束する.

3 構造の違いと固有値

前節までの議論から漸近挙動を支配しているのは、補題 3 で与えられた支配的固有値 λ_0 であることがわかった。また、補題 4 より、共通因数から得られる固有値によって決定付けられる解の挙動はすべて直線 l に沿っていることも分かる。これは、ネットワーク構造の違いに関わらず得られるものである。したがって、ネットワーク構造の違いは、解の漸近的な挙動よりも過渡的な挙動に影響を与え、また、その挙動は直線 l とは異なる方向の挙動であると考えられる。そこで我々は、共通因数 q_0 以外の因数 $q_i (i = 1, 2, 3)$ から得られる固有値についても解析を行った。

(I) 構造を $B = B_1$ で与えた場合。

命題 6.

q_1 からは実固有値は得られず、互いに共役な複素固有値 $\lambda_{\pm} = \alpha + i\beta$, ($\alpha < \lambda_0$, $\beta \neq 0$) が得られる。これら 2 つの固有値に属する固有関数は

$$e^{\alpha s} \begin{pmatrix} \cos \beta s \\ \cos(\beta - \frac{2}{3}\pi)s \\ \cos(\beta + \frac{2}{3}\pi)s \end{pmatrix}, \quad e^{\alpha s} \begin{pmatrix} \sin \beta s \\ \sin(\beta - \frac{2}{3}\pi)s \\ \sin(\beta + \frac{2}{3}\pi)s \end{pmatrix}$$

($-\tau \leq s \leq 0$) で与えられる。

固有関数は、線形化方程式(1)に対して l に垂直な面内で周期的に振動する解の成分を与えていることが分かる。すなわち、構造を $B = B_1$ で与えた場合の解挙動の特徴として、直線 l のまわりを振動的に直線 l へと近づき、その後定理 5 で述べたように l に沿って平衡点へと指数的に収束するのではないかということがあげられる。このことから、我々は、過渡的には振動を呈するのではないかと推測する。

(II) 構造を $B = B_2$ で与えた場合。

命題 7.

(i) $\frac{\tau q'(a)}{2e} e^{1+\frac{\tau}{e}} < 1$ のとき、 q_2 からは相異なる 2 つの実固有値 λ_2, λ'_2 が得られ $\lambda'_2 < \lambda_2 (< \lambda_0)$ 。それ以外の複素固有値を λ とすると $Re \lambda < \lambda'_2$ が成り立つ。 λ'_2, λ_2 も含めて固有関数は $e^{\lambda s} v$, ($-\tau \leq s \leq 0$) で与えられる。ここで v は v_0 に垂直な任意のベクトルである。

(ii) $\frac{\tau q'(a)}{2e} e^{1+\frac{\tau}{e}} > 1$ のとき q_2 からは実固有値は得られない。固有関数については命題 6 と同様。

構造を $B = B_2$ で与えた場合の特徴として、(i) 時間遅れが小さい場合; (ii) 時間遅れが大きい場合; の 2 通りが考えられる。時間遅れが小さい場合には、 l に垂直な面内の挙動を決定する固有値のうち λ_2, λ'_2 が実数値で与えられている。このことは、解は l へと振動することなく漸近し、その後 l に沿って平衡点へと収束することを示唆している。一方、時間遅れが大きい場合には、 $B = B_1$ の場合と同様である。つまり、過渡的振動は時間遅れが大きい場合に起こると推測できる。

(III) 構造を $B = B_3$ で与えた場合。

命題 8.

q_3 から得られる固有値は $\lambda = -\frac{1}{e}$ のみであり、固有関数は $e^{\lambda s} v$, ($-\tau \leq s \leq 0$) で与えられる。ここで v は v_0 に垂直な任意のベクトルである。

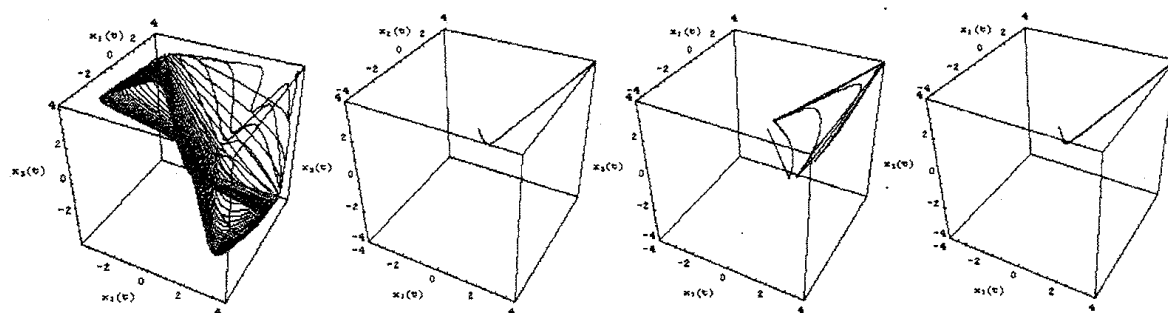
構造を $B = B_3$ で与えた場合の特徴としては、 $B = B_2$ の時間遅れが小さい場合と同様、 l へと振動することなく近づき、その後 l に沿って平衡点へと収束するのではないかということがあげられる。ただし、固有

値が実固有値ただひとつしか得られないことを考慮すると、 ℓ への近づき方は $B = B_2$ の場合に比べ、より単純ではないかと推察される。なお、この場合には過渡的な振動は起こらないと推測する。

4 数値計算

前節では、構造の違いが固有値に影響を与え、その結果を元に過渡的な振動の有無を推測した。しかし、固有値はあくまでも平衡点近傍での解の挙動を与えるものであり、平衡点から遠く離れたところで起こる過渡的な挙動については非線形の影響を無視することはできない。残念ながら、現段階で我々は非線形の影響を取り込んだ解析を行うまでに至っていない。したがって、前節での推測を裏付けるための数値計算例を紹介しよう。

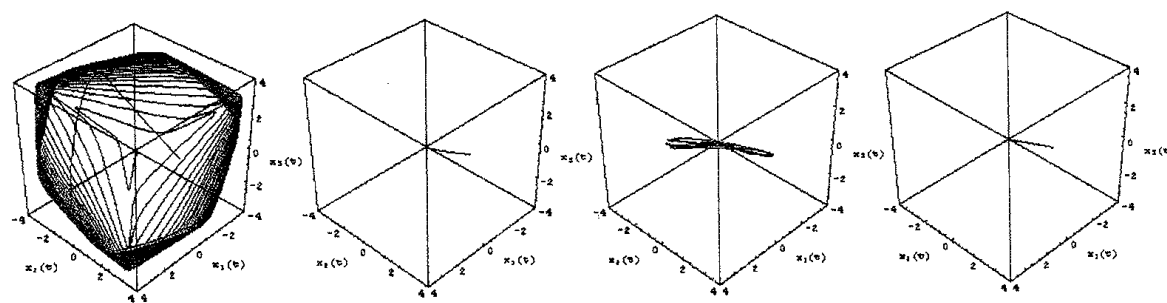
初期値を $(-1, 2, 0)$ 、 $\epsilon = 0.5$ 、 $W = 4$ として数値計算を行った。いずれの場合においても、解は平衡点 (a, a, a) 、 $(a = 3.9973\dots)$ に収束する様子が見て取れる。



(I) $B = B_1, \tau = 1$. (II-i) $B = B_2, \tau = 0.1$. (II-ii) $B = B_2, \tau = 4$. (III) $B = B_3, \tau = 0.1$.

図 2: 数値計算例

前節で推測したように、構造を $B = B_1$ で与えた場合 (図 2 (I)) や $B = B_2$ で時間遅れを大きくした場合 (図 2 (II-ii)) では、解は振動した後平衡点へと収束している。また、 $B = B_2$ で時間遅れが小さい場合 (図 2 (II-i)) や $B = B_3$ とした場合 (図 2 (III)) には速やかに直線 ℓ に漸近した後 ℓ に沿って平衡点へと収束している。なお、図 3 は、図 2 のグラフを直線 ℓ と垂直な面に射影したものである。



(I) $B = B_1, \tau = 1$. (II-i) $B = B_2, \tau = 0.1$. (II-ii) $B = B_2, \tau = 4$. (III) $B = B_3, \tau = 0.1$.

図 3: 図 2 の各グラフの直線 ℓ に垂直な面への射影

5 考察

構造の違いは、線形化方程式の固有値のうち、漸近的挙動を決定付ける支配的固有値には影響を与えず、それ以外の固有値に影響を与えることが分った。非線形方程式の過渡現象を線形化方程式の固有値で説明することに理論的な裏づけはない(少なくとも筆者らは知らない)が、数値計算結果とも照らし合わせると、これら構造の影響を受けた固有値と解の過渡的な挙動には関連があるように思われる。そして、 $B = B_1$ のときに観察される“長く続く過渡的な振動”は、 $B = B_2$ のときに時間遅れを大きくすることによって起きるが、 $B = B_3$ の場合や $B = B_2$ で時間遅れが小さい場合には起きないことが推測される。これは、情報の伝達量が少ない場合や、伝達に要する時間遅れが大きい場合には過渡的振動が起こるのではないかという直感的な事実とも一致しているのではないだろうか。

参考文献

- [1] P. Baldi and A. F. Atiya, “How delays affect neural dynamics and learning,” *IEEE Transactions On Neural Networks*, 5, 612–621, 1994.
- [2] T. Chu, Z. Zhang and Z. Wang, “A decomposition approach to analysis of competitive-cooperative neural networks with delay,” *Physics Letters A*, 312, 339–347, 2003.
- [3] C. M. Marcus and R. M. Westervelt, “Stability of analog neural networks with delay,” *Physical Review A*, 39, 347–359, 1994.
- [4] K. Pakdaman, C. P. Malta, C. Grotta-Ragazzo, O. Arino and J. -F. Vibert, “Transient oscillations in continuous-time excitatory ring neural networks with delay,” *Physical Review E*, 55, 3234–3248, 1997.
- [5] H. L. Smith, “Monotone Dynamical systems; An Introduction to the Theory of Competitive and Cooperative Systems, *Mathematical Surveys and Monographs*,” American Mathematical Society, Providence, 1995.