(2) 9-13

一种指数型动态神经网络的定性分析*

Qualitative Analysis of an Exponential Dynamic Neural Networks

傅 鹂　　裴小兵
Fu Li　　Pei Xiaobing

TP18

（重庆大学系统工程及应用数学系，重庆，630044；第一作者 35 岁，男，副教授，硕士）

摘　要　研究了由 T. D. Chiuch 和 R. M. Goodman 在 1988 年提出的指数型动态神经网络，给出了给定标准模式为该网络稳定点的一个简明的充分条件，并得到了这类稳定点的一步收敛半径。指出了在概率意义下系统不存在伪稳态。

关键词　神经网络；联想记忆；稳定点；伪稳态；收敛半径

指数型，定性分析

中国图书资料分类法分类号　TP302.7

ABSTRACT　The exponential dynamic neural network proposed by T. D. Chiuch and R. M. Goodman in 1988 is investigated. A concise sufficient condition for the given standard patterns to be stable is presented. A one-step convergence radius of such a stable state is obtained. It is also shown that there do not exist sporious stable states in the probabilistic sense.

KEYWORDS　neural networks; associative memory; stable state; spurious stable state; convergence radius

0　引　　言

Hopfield 网络用于联想记忆的存贮容量是较低的[1]，因此文献[2]提出了一种修正的 Hopfield 网络模型，通过 SNR 的途径可以证明该网络的存贮容量得到了明显改善[2,3]。这种修正的网络是在演化方程中引入了指数函数，它本质上是一种以含有指数运算为特征的动态系统，笔者称为指数型动态神经网络。笔者的研究进一步揭示了该网络的若干重要性质。在以下各节中分别给出了有关结果，即：待存贮的标准模式向量集为系统稳定点子集的一个简明的充分条件；标准模式稳定点的一步收敛半径；系统在概率意义下不存在伪稳态。本文所得出的理论结果为这类神经网络的设计及其联想记忆的实现提供了十分重要而直接的参考原则。

1　指数型动态神经网络模型

设神经网络的节点数为 n，待存贮的标准模式向量（以下简称标准模式）$x^1, \cdots, x^m$ 为 m 个 n 维双极性向量，即

* 收文日期　1996-03-15

$$x^k = (x_1^k, \cdots, x_n^k) \in \{-1, 1\}^n \qquad (k = 1, 2, \cdots, m, i = 1, 2, \cdots, n) \tag{1}$$

设 $x(t) \in \{-1, 1\}^n$, $x(t+1) \in \{-1, 1\}^n$ 为网络在 t 时刻和 $t+1$ 时刻的节点状态，则指数型神经网络模型的演化方程为：

$$x(t+1) = \mathrm{sgn}\left(\sum_{k=1}^{m} x^k \cdot a^{x^k \cdot x(t)}\right), \qquad (a > 1) \tag{2}$$

其中

$$\mathrm{sgn}(u) = \begin{cases} 1, & u \geqslant 0 \\ -1, & u < 0 \end{cases} \tag{3}$$

为符号函数，而

$$x^k \cdot x(t) = \sum_{i=1}^{n} x_i^k \cdot x_i(t) \tag{4}$$

表示内积。以上模型可对双极性模式向量直接处理。对一般的模式向量(向量的分量为实数)，考虑到实际数据的位数总是限的，因此可以通过适当的编码化为双极性模式，不存在原则上的困难。此外，这一模型的参数已由标准模式直接设定，故不须复杂的学习或训练过程。

2 标准式与稳定点的关系

根据动态神经网络联想记忆的原则，希望所有待存贮的标准式向量均对应于网络的稳定点。下列定理给出了一个简明的充分条件。

定理 1 在演化方程(2)，若 $a > \sqrt{m-1}$ 则 $\{x^1, x^2, \cdots, x^m\}$ 均为系统的稳定点。

证明： 由(2)得相应分量形式：

$$x_i(t+1) = \mathrm{sgn}\left(\sum_{k=1}^{m} x_i^k \cdot a^{x^k \cdot x(t)}\right), \qquad (i = 1, 2, \cdots, n) \tag{5}$$

令 $x(t) = x^h (h = 1, 2, \cdots, m)$，则：

$$x_i(t+1) = \mathrm{sgn}\left(\sum_{k=1}^{m} x_i^k \cdot a^{x^k \cdot x^h}\right) = \mathrm{sgn}\left(x_i^h \cdot a^n + \sum_{\substack{k=1 \\ k \neq h}}^{m} x_i^k \cdot a^{x^k \cdot x^h}\right) \tag{6}$$

因为 $x^k \neq x^h, k \neq h, k = 1, 2, \cdots, m$，故 $x^k \cdot x^h \leqslant n - 2$，所以当 $a^n > (m-1)a^{n-2}$ 即 $a > \sqrt{m-1}$ 时有：

$$x_i(t+1) = x_i^h, \qquad i = 1, 2, \cdots, n \tag{7}$$

可见，当 $a > \sqrt{m-1}$ 时，若 $x(t) = x^h$ 为某个标准模式，则 $x(t+1) = x^h$，即每个标准模式均为系统的稳定点。证毕。

此定理表明只要演化方程(2)中的参数 a 取得足够大，便可存贮所有预先给定的模式向量，而不需要标准模式向量之间必须线性无关等前提条件。

3 收敛半径

下面讨论标准模式的吸引域问题。由于针对双极性模式，故采用汉明距离作度量。

引理 若 x, y, z 均为 n 维双极性向量，H 表示汉明距离(下同)，则有：

$$H(x, z) \leqslant H(x, y) + H(y, z) \tag{8}$$

$$x \cdot y = n - 2H(x,y) \tag{9}$$

证明：(略)

以下定理给出了模型(2) 标准模式稳定点的一步收敛半径。

定理 2　设 $x(t)$ 为输入模式，若定理 1 条件成立即 $a > \sqrt{m-1}$，且存在 $x^s \in \{x^1, \cdots, x^m\}$ 为标准模式集，使 $H(x^s, x(t)) < H_s/2$，$H_s = \min\{H(x^s, x^k), k = 1,2,\cdots,m; k \neq s\}$，则 $x(t+1) = \mathrm{sgn}(\sum_{k=1}^{m} x^k a^{x^k \cdot x(t)}) = x_s$.

证明：由(5) 式有：

$$x_i(t+1) = \mathrm{sgn}(x_i^s \cdot a^{x^s \cdot x(t)} + \sum_{\substack{k=1 \\ k \neq s}}^{m} x_i^k \cdot a^{x^k \cdot x(t)}) \qquad (i = 1,2,\cdots,m) \tag{10}$$

由(9) 知

$$x^k \cdot x(t) = n - 2H(x^k, x(t)), \qquad (k = 1,2,\cdots,m) \tag{11}$$

所以

$$x_i(t+1) = \mathrm{sgn}(x_i^s \cdot a^{(n-2H(x^s,x(t)))} + \sum_{\substack{k=1 \\ k \neq s}}^{m} x_i^k \cdot a^{(n-2H(x^k,x(t)))}) \qquad (i = 1,2,\cdots,n) \tag{12}$$

由(8) 得：

$$H_s \leqslant H(x^s, x(t)) + H(x^k, x(t)) \tag{13}$$

可见欲使 $x(t+1) = x^s$，只需 $a^{(n-2H(x^s,x(t)))} > (m-1)a^{(n-2H_s+2H(x^s,x(t)))}$ 或等价地

$$a^{[H_s - 2H(x^s,x(t))]} > \sqrt{m-1} \tag{14}$$

由本定理的假设条件 $a > \sqrt{m-1}$ 及 $H(x^s, x(t)) < H_s/2$ 立即知上式成立。证毕。

此定理表明，对任何一输入模式 $x(t)$，如果 $x(t)$ 落在以某个标准模式 x^s 为中心以 $H_s/2$ 为半径的邻域内，则 $x(t)$ 将进一步收敛到 x^s.

4　收敛性及伪稳态问题

上节证明了，只要初始状态与某个标准模式的距离小于收敛半径($H_s/2$)，则必一步收敛到该标准模式。本节将证明，在概率意义下，系统从任意初始状态都将收敛到某个标准模式，换言之，在概率意义下系统不仅收敛而且不存在伪稳态。

定义 1　设 x 为 n 维双极性模式，将 x 扩充为 $p(p > n)$ 维的双极性模式 $\tilde{x}$，其中 $x_i = \tilde{x}_i$，$i = 1,2,\cdots,n$；$\tilde{x}_i = 1$，$i = n+1,\cdots,p$. 则称 $\tilde{x}$ 为 x 的增广向量。

由定义易知

$$H(x,y) = H(\tilde{x}, \tilde{y}) \tag{15}$$

定义 2　系统(2) 的一个稳态若不是标准模式，则称此状态为伪稳态。

伪稳态是不期望存贮的模式，也是不应联想记忆的模式。因此，如何才能避免这些伪稳定态的出现，是研究联想记忆必须要解决的问题。

由于实际问题中的标准模式可以是任意的，因此，假定标准模式$\{x^1,\cdots,x^m\}$为独立的均匀分布的双极性随机向量才是合理的，以下的结果是以这个假定为前提的。

定理3　设标准模式$\{x^1,\cdots,x^m\}$是相互独立服从均匀分布的双极性随机向量，则任意初态$x(t)$收敛到某个标准模式x^s的概率为1.

证明：设$\{x^k,k=1,2,m\}$，$x(t)$，$x(t+1)$的增广向量分别为$\{\tilde{x}^k,k=1,2,\cdots,m\}$，$\tilde{x}(t)$，$\tilde{x}(t+1)$，且增广向量的维数为$p$. 则

$$\tilde{x}_i(t+1)=\mathrm{sgn}(\sum_{k=1}^{m}\tilde{x}^k\cdot a^{\tilde{x}^k\cdot\tilde{x}(t)})\tag{16}$$

或等价地

$$\tilde{x}_i(t+1)=\mathrm{sgn}(\sum_{k=1}^{m}\tilde{x}_i^k\cdot a^{\tilde{x}^k\cdot\tilde{x}(t)}),\qquad(i=1,2,\cdots,p)\tag{17}$$

对某个(增广)标准模式向量$\tilde{x}^s$上式可写成

$$\tilde{x}_i(t+1)=\mathrm{sgn}(\tilde{x}_i^s\cdot a^{\tilde{x}^s\cdot\tilde{x}(t)})+\sum_{\substack{k=1\\k\neq s}}^{m}\tilde{x}_i^k\cdot a^{\tilde{x}^k\cdot\tilde{x}(t)})\tag{18}$$

令$\eta_k=\tilde{x}_i^k a^{\tilde{x}^k\cdot\tilde{x}(t)}$由定理条件知$\{\eta_k\}$是独立同分布随机变量，故由文献[2]有下列两式：

$$P_r\{\eta_1=a^{p-2k}\}=2^{-p}C_{p-1}^k\tag{19}$$

$$P_r\{\eta_1=-a^{p-2k-2}\}=2^{-p}C_{p-1}^k,H(\tilde{x}^1,\tilde{x}(t))=k,\qquad k=0,1,\cdots,P-1\tag{20}$$

其中P_r表示概率，由此得η_1的期望为：

$$E[\eta_1]=a^p\cdot2^{-p}(1-a^2)(1+a^{-2})^{p-1}\tag{21}$$

同时还可求得$E[\eta_1^2]=[(1+a^{-4})\cdot(a^2/2)]^p$，所以方差

$$D[\eta_1^2]=E[\eta_1^2]-E^2[\eta_1]=(a/2)^{2p}[(2+2a^{-1})^p-(1+a^{-2})^{2(p-1)}(1-a^{-2})^2]\tag{22}$$

若$\tilde{x}_i^s=1$，则：

$$P_r\{\tilde{x}_i(t+1)=-1\}=P_r\{a^{\tilde{x}\cdot\tilde{x}(t)}+\sum_{\substack{k=1\\k\neq s}}^{m}\tilde{x}_i^k\cdot a^{\tilde{x}^k\cdot\tilde{x}(t)}<0\}\tag{23}$$

若$x_i^s=-1$，则：

$$P_r\{\tilde{x}_i(t+1)\}=P_r\left\{\tilde{x}_i^s\cdot a^{\tilde{x}^s\tilde{x}(t)}+\sum_{\substack{k=1\\k\neq s}}^{m}a^{\tilde{x}^k\cdot\tilde{x}(t)}\tilde{x}_i^k\geqslant0\right\}=P_r\left\{-a^{\tilde{x}^s\cdot\tilde{x}(t)}+\sum_{\substack{k=1\\k\neq s}}^{m}a^{\tilde{x}^k\cdot\tilde{x}(t)}\tilde{x}_i^k\geqslant0\right\}\tag{24}$$

设$P_0=P_r\{\tilde{x}_i^s\neq\tilde{x}_{i+1}(t+1)\}$，则：

$$\begin{aligned}P_0&=P_r\{\tilde{x}_i(t+1)=1\mid\tilde{x}_i^s=-1\}+P_r\{\tilde{x}_i(t+1)=-1\mid\tilde{x}_i^s=1\}\\&=\left\{\left|\sum_{\substack{k=1\\k\neq s}}^{m}a^{\tilde{x}^k\cdot\tilde{x}(t)}\cdot\tilde{x}_i^k\right|\geqslant a^{\tilde{x}^s\cdot\tilde{x}(t)}\right\}-P_r\left\{\sum_{\substack{k=1\\k\neq s}}^{m}a^{\tilde{x}^k\cdot\tilde{x}(t)}\cdot\tilde{x}_i^k\geqslant-a^{\tilde{x}^s\cdot\tilde{x}(t)}\right\}\end{aligned}$$

$$\leqslant \frac{m-1)D(\eta_1)}{a^{2(\tilde{x}^s \cdot \tilde{x}(t))}} - P_r\left\{ \left| \sum_{\substack{k=1 \\ k \neq s}}^{m} a^{\tilde{x}^k \cdot \tilde{x}(t)} \cdot \tilde{x}_i^k \right| \geqslant a^{\tilde{x}^s \cdot \tilde{x}(t)} \right\}$$

$$\leqslant \frac{m-1)D(\eta_1)}{a^{2(\tilde{x}^s \cdot \tilde{x}(t))}} \tag{25}$$

所以

$$P_r\{x(t+1) = x^s\} = P_r\{\tilde{x}(t+1) = \tilde{x}^s\} = 1 - P_r\{\tilde{x}(t+1) \neq \tilde{x}^s\}$$

$$= 1 - P_r\{\bigcup_{i=1}^{p}[\tilde{x}_i(t+1) \neq \tilde{x}_i^s]\} \geqslant 1 - \sum_{i=1}^{p} P_r\{\tilde{x}_i(t+1) \neq \tilde{x}_i^s\}$$

$$\geqslant 1 - p\,\frac{(m-1)D(\eta_1)}{a^{2p}} \cdot a^{4H(\tilde{x}^s \cdot \tilde{x}(t))}$$

$$= 1 - \frac{p(m-1)}{4^p}\left[(2+2a^{-1})^p - (1+a^{-2})^{2(p-1)}(1-a^{-2})^2\right] \cdot a^{4H(\tilde{x}^s \cdot \tilde{x}(t))}$$

$$\geqslant 1 - \frac{p(m-1)}{2^p}(1+a^{-1})^p - a^{4H(\tilde{x}^s \cdot \tilde{x}(t))} \geqslant 1 - \frac{p(m-1)}{2^p}(1+a^{-1})^p - a^{4H(x^s \cdot x(t))}$$

$$\geqslant 1 - \frac{p(m-1)}{2^p}(1+a^{-1})^p - a^{4n} \xrightarrow{(p\to\infty)} 1$$

由于事实上 $P_r\{x(t+1) = x^s\}$ 与 p 无关，故

$$P_r\{x(t+1) = x^s\} = 1 \tag{26}$$

即任意初态按概率 1（一步）收敛到某个标准模式，或等价地，存在伪稳态的概率为 0. 证毕。

5 结　语

本文对文献[2]提出的指数型神经网络进行了深入的分析，揭示了其若干重要性质，有关的理论结果为这类网络的设计及其联想记忆的实现提供了非常简洁的参考准则。为了进一步提高这类网络的性能而做的推广研究正在进行之中。

参 考 文 献

1　Meliece R J, Posner E C, Rodemich E R, Venkatesh S S. The capacity of the Hopfield associative memory. IEEE Trans. on Information Theory. 1987, 33(4): 461～482

2　Chiueh T D, Goodman R M. High-capacity exponential associative memory. Proc. IFCNN, Vol. I. IEEE Inc., New York: 1988. 153～160

3　Wang C C, Don H S. An analysis of high-capacity discrete exponential BAM. IEEE Trans. on Neural Networks. 1995, 6(2): 492～596

4　Hecht-Niecsen R. Neurocompution. IEEE Spectrum. March, 1988. 36～41

5　焦李成．神经网络系统理论．西安：西安电子科技大学出版社．1992. 52～88

6　杨行峻，郑君里．人工神经网络．北京：高等教育出版社．1992. 20～55