

基于一类新型激活函数的递归神经网络的多稳定性分析

闫维昕, 刘洋, 王震

(山东科技大学 数学与系统科学学院, 山东 青岛 266590)

摘要:本研究基于一类新型连续锯齿型激活函数研究了递归神经网络(Hopfield神经网络)的多稳定性。首先,通过区间剖分法、Brouwer 不动点定理证明了基于该激活函数的 n 维神经网络模型至少具有 7^n 个平衡点,并运用对角占优矩阵理论、局部稳定性判定定理等方法证明了其中 4^n 个平衡点是局部指数稳定的,剩余的平衡点是不稳定的。其次,通过增加锯齿型激活函数的峰值点将激活函数推广到更一般的情况,得到了 n 维神经网络在含有 k 个峰值点的连续锯齿型激活函数中至少具有 $(2k+1)^n$ 个平衡点,其中 $(k+1)^n$ 个平衡点为局部稳定的。本研究设计的激活函数相较于现有的一些激活函数会产生更多的稳定平衡点,并且在增加峰值点的过程中不会增加神经网络的计算复杂度。最后,通过两个具体的数值算例验证了本研究结果的有效性。

关键词: Hopfield 神经网络; 多稳定性; 平衡点; 连续锯齿型激活函数

中图分类号: N941.3; N941.7

文献标志码: A

Multistability analysis of recurrent neural networks with a new type of activation functions

YAN Weixin, LIU Yang, WANG Zhen

(College of Mathematics and Systems Science, Shandong University of Science and Technology, Qingdao 266590, China)

Abstract: In this paper, the multistability of Hopfield neural network was studied based on a new type of activation function, the continuous sawtooth activation function. Firstly, the n -neural network model based on this activation function was proved to have at least 7^n equilibrium points through interval partition method and Brouwer's fixed point theorem. By using the diagonally dominant matrix theory and local stability judgment theorem, 4^n equilibrium points were proved to be locally exponentially stable while the remaining equilibrium points to be unstable. Secondly, the activation function was generalized to a more general situation by adding the peak points of the sawtooth activation function. The n -neural network was found to have at least $(2k+1)^n$ equilibrium points in the continuous sawtooth activation function with k peak points, of which $(k+1)^n$ equilibrium points are local stable. Compared with the existing activation functions, the activation function designed in this paper can generate more stable equilibrium points and will not increase the computational complexity of the neural network in the process of increasing the number of peak points. Finally, two numerical examples were presented to demonstrate the validity of the results.

Key words: Hopfield neural network; multistability; equilibrium point; continuous sawtooth-type activation function

近年来,神经网络在图像处理、模式识别、联想记忆等领域的广泛应用^[1-3],吸引了众多学者的研究兴趣。神经网络系统在激活函数作用下会呈现不同的动力学特性,而对网络的动力学行为研究是神经网络在各领域中广泛应用的前提,其中稳定性分析成为神经网络设计和应用必不可少的条件^[4]。神经网络的稳定性分

收稿日期:2022-12-02

基金项目:国家自然科学基金项目(62173214);山东省自然科学基金项目(ZR2021MF003)。

作者简介:闫维昕(2000—),女,山西大同人,硕士研究生,主要从事神经网络稳定性研究。

王震(1977—),男,山东泰安人,教授,博士生导师,主要从事非线性系统理论研究,本文通信作者。

E-mail: wangzhen@sdust.edu.cn

为单稳定性和多稳定性问题。对于单稳定性的研究已经有了一些成果^[5-6],但单稳定性在复杂问题的计算中提供的计算能力稍显不足。神经网络中局部稳定平衡点的数目与其存储能力相对应,拥有多个稳定平衡点意味着拥有更大的存储容量。如在联想记忆的应用^[7]中,记忆模态通常以稳定平衡点的形式存储起来,一个稳定平衡点对应于一种记忆模态,神经网络的解收敛到一个稳定平衡点的过程代表一个联想记忆的过程。因此,具有联想记忆功能的神经网络拥有的稳定平衡点越多,可以储存的记忆模态越多,储存容量越大。

在神经网络多稳定性分析中,多个平衡点的存在性和稳定性是目前主要研究方向。常见的用于分析神经网络多稳定性的激活函数有单调递增的 Sigmoid 函数^[8]和饱和函数^[9]。与单调函数相较,采用非单调函数作为激活函数的神经网络会有更多的平衡点。文献^[10]使用一种新型激活函数——墨西哥帽型函数来讨论神经网络的多稳定性,该函数为连续非单调线性函数,在一些假设条件下存在 3^n 个平衡点,其中 2^n 个平衡点是局部稳定的;文献^[11]提出一类高斯小波型激活函数,具有该函数的神经网络存在 5^n 个平衡点,其中 3^n 个平衡点是局部稳定的。显然,具有高斯小波型激活函数的神经网络比具有饱和函数、Sigmoid 函数以及墨西哥帽型激活函数的神经网络具有更多的平衡点和稳定点。

本研究基于连续锯齿型激活函数研究了 Hopfield 神经网络的多稳定性问题,主要贡献为:首先,本研究设计的连续锯齿型激活函数相较于现有的一些激活函数能产生更多的稳定平衡点,意味着神经网络的储存容量得到提升,在联想记忆的应用中可以储存更多的记忆模式;其次,增加该激活函数峰值点的过程不会增加神经网络的计算复杂度,更易于将该激活函数应用在实际中。

1 问题的提出

1.1 激活函数的设计

为进一步增加神经网络中稳定点的数目,本研究设计了一种新型连续锯齿型激活函数,该函数可以通过增加峰值点的方式有效扩展平衡点的数目。激活函数定义如下。

$$f_i(x) = \begin{cases} u_i, & -\infty < x < m_{i,1}; \\ l_{i,1}x + c_{i,1}, & m_{i,1} \leq x < n_{i,1}; \\ l_{i,2}x + c_{i,2}, & n_{i,1} \leq x < m_{i,2}; \\ l_{i,3}x + c_{i,3}, & m_{i,2} \leq x < n_{i,2}; \\ l_{i,4}x + c_{i,4}, & n_{i,2} \leq x < m_{i,3}; \\ l_{i,5}x + c_{i,5}, & m_{i,3} \leq x < n_{i,3}; \\ v_i, & n_{i,3} \leq x < +\infty. \end{cases} \quad (1)$$

式中: $m_{i,1}, n_{i,1}, m_{i,2}, n_{i,2}, m_{i,3}, n_{i,3}, l_{i,1}, l_{i,2}, l_{i,3}, l_{i,4}, l_{i,5}, c_{i,1}, c_{i,2}, c_{i,3}, c_{i,4}, c_{i,5}, u_i, v_i$ 均为常数,并且满足 $-\infty < m_{i,1} < n_{i,1} < m_{i,2} < n_{i,2} < m_{i,3} < n_{i,3} < +\infty, l_{i,1} > 0, l_{i,2} < 0, l_{i,3} > 0, l_{i,4} < 0, l_{i,5} > 0, u_i = f_i(m_{i,1}) = f_i(m_{i,2}) = f_i(m_{i,3}), v_i = f_i(n_{i,1}) = f_i(n_{i,2}) = f_i(n_{i,3}), f_i(m_{i,1}) < f_i(n_{i,1}), i = 1, 2, \dots, n$ 。该激活函数中存在 3 个峰值点,分别为 $(n_{i,1}, f_i(n_{i,1})), (n_{i,2}, f_i(n_{i,2})), (n_{i,3}, f_i(n_{i,3}))$ 。

连续锯齿型激活函数为连续非单调的激活函数,锯齿波是电路中常见的波形之一,如示波管电路中的扫描电压和偏转电流,经过运算放大器的饱和和处理可以应用于显像管显示和故障检测等方面。相比于双曲正切函数与 sigmoid 函数,基于连续锯齿型激活函数的神经网络可以产生更多数目的稳定平衡点,即神经网络具有更大的储存容量。

1.2 模型介绍

本研究考虑如下神经网络模型的动力学方程:

$$\frac{dx_i(t)}{dt} = -d_i x_i(t) + \sum_{j=1}^n a_{ij} f_j(x_j(t)) + J_i, \quad i = 1, 2, \dots, n. \quad (2)$$

用 $\mathbf{x} = (x_1, x_2, \dots, x_n)^T \in \mathbf{R}^n$ 表示状态向量; $\mathbf{D} = \text{diag}(d_1, d_2, \dots, d_n)$ 为正定矩阵, $d_i > 0$; $\mathbf{A} = (a_{ij})$ 为连接权矩阵; $f(\mathbf{x}) = (f_1(x_1), f_2(x_2), \dots, f_n(x_n))^T$ 表示神经元的非线性激活函数; $\mathbf{J} = (J_1, J_2, \dots, J_n)^T$ 表示

系统的输入向量。

根据激活函数(式(1))的性质,将实数域 \mathbf{R} 分解为下述区间:

$$\begin{aligned} (-\infty, m_{i,1}) &= (-\infty, m_{i,1})^1 \times (m_{i,1}, n_{i,1})^0 \times (n_{i,1}, m_{i,2})^0 \times (m_{i,2}, n_{i,2})^0 \times (n_{i,2}, m_{i,3})^0 \times (m_{i,3}, n_{i,3})^0 \times (n_{i,3}, +\infty)^0, \\ (m_{i,1}, n_{i,1}) &= (-\infty, m_{i,1})^0 \times (m_{i,1}, n_{i,1})^1 \times (n_{i,1}, m_{i,2})^0 \times (m_{i,2}, n_{i,2})^0 \times (n_{i,2}, m_{i,3})^0 \times (m_{i,3}, n_{i,3})^0 \times (n_{i,3}, +\infty)^0, \\ (n_{i,1}, m_{i,2}) &= (-\infty, m_{i,1})^0 \times (m_{i,1}, n_{i,1})^0 \times (n_{i,1}, m_{i,2})^1 \times (m_{i,2}, n_{i,2})^0 \times (n_{i,2}, m_{i,3})^0 \times (m_{i,3}, n_{i,3})^0 \times (n_{i,3}, +\infty)^0, \\ (m_{i,2}, n_{i,2}) &= (-\infty, m_{i,1})^0 \times (m_{i,1}, n_{i,1})^0 \times (n_{i,1}, m_{i,2})^0 \times (m_{i,2}, n_{i,2})^1 \times (n_{i,2}, m_{i,3})^0 \times (m_{i,3}, n_{i,3})^0 \times (n_{i,3}, +\infty)^0, \\ (n_{i,2}, m_{i,3}) &= (-\infty, m_{i,1})^0 \times (m_{i,1}, n_{i,1})^0 \times (n_{i,1}, m_{i,2})^0 \times (m_{i,2}, n_{i,2})^0 \times (n_{i,2}, m_{i,3})^1 \times (m_{i,3}, n_{i,3})^0 \times (n_{i,3}, +\infty)^0, \\ (m_{i,3}, n_{i,3}) &= (-\infty, m_{i,1})^0 \times (m_{i,1}, n_{i,1})^0 \times (n_{i,1}, m_{i,2})^0 \times (m_{i,2}, n_{i,2})^0 \times (n_{i,2}, m_{i,3})^0 \times (m_{i,3}, n_{i,3})^1 \times (n_{i,3}, +\infty)^0, \\ (n_{i,3}, +\infty) &= (-\infty, m_{i,1})^0 \times (m_{i,1}, n_{i,1})^0 \times (n_{i,1}, m_{i,2})^0 \times (m_{i,2}, n_{i,2})^0 \times (n_{i,2}, m_{i,3})^0 \times (m_{i,3}, n_{i,3})^0 \times (n_{i,3}, +\infty)^1. \end{aligned}$$

并且将集合 Ω_n 定义为:

$$\begin{aligned} \Omega_n &= \left\{ \prod_{i=1}^n (-\infty, m_{i,1})^{\delta_1^{(i)}} \times [m_{i,1}, n_{i,1}]^{\delta_2^{(i)}} \times (n_{i,1}, m_{i,2})^{\delta_3^{(i)}} \times [m_{i,2}, n_{i,2}]^{\delta_4^{(i)}} \times (n_{i,2}, m_{i,3})^{\delta_5^{(i)}} \times [m_{i,3}, n_{i,3}]^{\delta_6^{(i)}} \times \right. \\ &\left. (n_{i,3}, +\infty)^{\delta_7^{(i)}} \right\}, (\delta_1^{(i)}, \delta_2^{(i)}, \delta_3^{(i)}, \delta_4^{(i)}, \delta_5^{(i)}, \delta_6^{(i)}, \delta_7^{(i)}) = (1, 0, 0, 0, 0, 0, 0) \text{ 或 } (0, 1, 0, 0, 0, 0, 0) \text{ 或 } (0, 0, 1, 0, 0, 0, 0) \\ &\text{或 } (0, 0, 0, 1, 0, 0, 0) \text{ 或 } (0, 0, 0, 0, 1, 0, 0) \text{ 或 } (0, 0, 0, 0, 0, 1, 0) \text{ 或 } (0, 0, 0, 0, 0, 0, 1), i = 1, 2, \dots, n, \text{ 可以得到} \\ &\text{到集合 } \Omega_n \text{ 由 } 7^n \text{ 个区域组成。} \end{aligned}$$

设 ϵ 为一个足够小的正数,且满足:

$$0 < \epsilon \ll \min_{1 \leq i \leq n} \left\{ \left| \frac{1}{m_{i,1}} \right|, \left| \frac{1}{n_{i,3}} \right|, \frac{d_i}{\sum_{j=1}^n |a_{ij}| \mu_j + |J_i|} \right\}. \quad (3)$$

式中: $\mu_j = \max\{|u_j|, |v_j|\}, j = 1, 2, \dots, n$ 。在此定义下再给出集合 Ω_ϵ 的定义:

$$\begin{aligned} \Omega_\epsilon &= \left\{ \prod_{i=1}^n \left[\frac{-1}{\epsilon}, m_{i,1} - \epsilon \right]^{\delta_1^{(i)}} \times [m_{i,1} + \epsilon, n_{i,1} - \epsilon]^{\delta_2^{(i)}} \times [n_{i,1} + \epsilon, m_{i,2} - \epsilon]^{\delta_3^{(i)}} \times [m_{i,2} + \epsilon, n_{i,2} - \epsilon]^{\delta_4^{(i)}} \times \right. \\ &\left. [n_{i,2} + \epsilon, m_{i,3} - \epsilon]^{\delta_5^{(i)}} \times [m_{i,3} + \epsilon, n_{i,3} - \epsilon]^{\delta_6^{(i)}} \times \left[n_{i,3} + \epsilon, \frac{1}{\epsilon} \right]^{\delta_7^{(i)}} \right\}, (\delta_1^{(i)}, \delta_2^{(i)}, \delta_3^{(i)}, \delta_4^{(i)}, \delta_5^{(i)}, \delta_6^{(i)}, \delta_7^{(i)}) = \\ &(1, 0, 0, 0, 0, 0, 0) \text{ 或 } (0, 1, 0, 0, 0, 0, 0) \text{ 或 } (0, 0, 1, 0, 0, 0, 0) \text{ 或 } (0, 0, 0, 1, 0, 0, 0) \text{ 或 } (0, 0, 0, 0, 1, 0, 0) \\ &\text{或 } (0, 0, 0, 0, 0, 1, 0) \text{ 或 } (0, 0, 0, 0, 0, 0, 1), i = 1, 2, \dots, n. \text{ 对任意集合 } \Omega(m) \in \Omega_\epsilon, m \in \{1, 2, \dots, 7^n\}, \Omega(m) \text{ 都为} \\ &\text{一个有界闭集。} \end{aligned}$$

2 主要分析证明

2.1 平衡点的存在性分析

以下对基于连续锯齿型激活函数(式(1))的 Hopfield 神经网络(式(2))平衡点的存在性进行分析,即分析其平衡点的个数。首先给出下述定理。

定理 1 在集合 Ω_ϵ 中,若满足两个条件(式(4)、式(5)),则基于连续锯齿型激活函数(式(1))的递归神经网络(式(2))至少存在 7^n 个平衡点。

$$-d_i m_{i,1} + a_{ii} u_i + \sum_{j=1, j \neq i}^n \max\{a_{ij} u_j, a_{ij} v_j\} + J_i < 0, \quad (4)$$

$$-d_i n_{i,3} + a_{ii} v_i + \sum_{j=1, j \neq i}^n \min\{a_{ij} u_j, a_{ij} v_j\} + J_i > 0. \quad (5)$$

式中: $i = 1, 2, \dots, n$ 。

证明: 首先,从集合 Ω_ϵ 中任意选取一个区域,并将其表示为:

$$\tilde{\Omega}_\epsilon = \prod_{i \in N_1} \left[-\frac{1}{\epsilon}, m_{i,1} - \epsilon \right] \times \prod_{i \in N_2} [m_{i,1} + \epsilon, n_{i,1} - \epsilon] \times \prod_{i \in N_3} [n_{i,1} + \epsilon, m_{i,2} - \epsilon] \times \prod_{i \in N_4} [m_{i,2} + \epsilon, n_{i,2} - \epsilon] \times$$

$$\prod_{i \in N_5} [n_{i,2} + \varepsilon, m_{i,3} - \varepsilon] \times \prod_{i \in N_6} [m_{i,3} + \varepsilon, n_{i,3} - \varepsilon] \times \prod_{i \in N_7} \left[n_{i,3} + \varepsilon, \frac{1}{\varepsilon} \right] \subset \Omega_\varepsilon.$$

式中: N_i 为集合 $\{1, 2, \dots, n\}$ 的子集, 并且 $N_i \cap N_j = \emptyset (i \neq j, i, j = 1, 2, \dots, 7)$, $N_1 \cup N_2 \cup N_3 \cup N_4 \cup N_5 \cup N_6 \cup N_7 = \{1, 2, \dots, n\}$ 。接下来证明, 具有连续锯齿型激活函数(式(1))的递归神经网络(式(2))在区域 $\tilde{\Omega}_\varepsilon$ 中至少存在一个平衡点。

在区域 $\tilde{\Omega}_\varepsilon$ 中任意选取一点 $(x_1, x_2, \dots, x_n)^T$, 将分量 x_i 取出, 固定其余分量 $x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_n$ 。定义函数 $F_i(\xi) = -d_i \xi + a_{ii} f_i(\xi) + \sum_{j=1, j \neq i}^n a_{ij} f_j(x_j) + J_i$, $\xi \in (-\infty, \infty)$, 根据区域的划分情况, 共有以下 7 种情况需要进行讨论。

情况 1 当 $i \in N_1$ 时, 从设定的已知条件中可以知道 $u_i = f_i(m_{i,1} - \varepsilon)$, 并且 $u_j \leq f_j(x_j) \leq v_j$, 由 ε 的定义(式(3))与定理 1 中的条件(4), 可以得到不等式:

$$F_i\left(-\frac{1}{\varepsilon}\right) \geq \frac{d_i}{\varepsilon} - \sum_{j=1}^n |a_{ij}| \mu_j - |J_i| > 0,$$

$$F_i(m_{i,1} - \varepsilon) \leq -d_i m_{i,1} + d_i \varepsilon + a_{ii} u_i + \sum_{j=1, j \neq i}^n \max\{a_{ij} u_j, a_{ij} v_j\} + J_i \leq 0.$$

再由函数 $F_i(\xi)$ 的连续性, 存在一点 $\bar{x}_i \in \left(-\frac{1}{\varepsilon}, m_{i,1} - \varepsilon\right]$ 使得 $F_i(\bar{x}_i) = 0$ 。

情况 2 当 $i \in N_2$ 时, 已知 $f_i(m_{i,1}) = u_i$, 并且 $f_i(n_{i,1}) = f_i(n_{i,3}) = v_i, n_{i,1} < n_{i,3}$, 由定理 1 中的式(4)和式(5), 可以得到:

$$F_i(m_{i,1} + \varepsilon) \leq -d_i m_{i,1} - d_i \varepsilon + a_{ii} f(m_{i,1} + \varepsilon) + \sum_{j=1, j \neq i}^n \max\{a_{ij} u_j, a_{ij} v_j\} + J_i < 0,$$

$$F_i(n_{i,1} - \varepsilon) > -d_i n_{i,3} + d_i \varepsilon + a_{ii} v_i + \sum_{j=1, j \neq i}^n \min\{a_{ij} u_j, a_{ij} v_j\} + J_i > 0.$$

由函数 $F_i(\xi)$ 的连续性可知, 存在一点 $\bar{x}_i \in (m_{i,1} + \varepsilon, n_{i,1} - \varepsilon)$ 使得 $F_i(\bar{x}_i) = 0$ 。

情况 3 当 $i \in N_3$ 时, 已知 $f_i(m_{i,2}) = u_i$, 并且 $f_i(n_{i,1}) = f_i(n_{i,3}) = v_i, m_{i,1} < m_{i,2}, n_{i,1} < n_{i,3}$ 。由定理 1 中的式(4)和式(5), 可以得到不等式:

$$F_i(n_{i,1} + \varepsilon) > -d_i n_{i,3} - d_i \varepsilon + a_{ii} f_i(n_{i,1} + \varepsilon) + \sum_{j=1, j \neq i}^n \min\{a_{ij} u_j, a_{ij} v_j\} + J_i \geq 0,$$

$$F_i(m_{i,2} - \varepsilon) < -d_i m_{i,1} + d_i \varepsilon + a_{ii} f_i(m_{i,2} - \varepsilon) + \sum_{j=1, j \neq i}^n \max\{a_{ij} u_j, a_{ij} v_j\} + J_i \leq 0.$$

由函数 $F_i(\xi)$ 的连续性可知, 存在一点 $\bar{x}_i \in [n_{i,1} + \varepsilon, m_{i,2} - \varepsilon]$, 使得 $F_i(\bar{x}_i) = 0$ 。

情况 4 当 $i \in N_4$ 时, 已知 $f_i(m_{i,2}) = u_i, m_{i,1} < m_{i,2}$, 并且 $f_i(n_{i,2}) = v_i$, 由定理 1 中的式(4)和式(5), 可以得到不等式:

$$F_i(m_{i,2} + \varepsilon) < -d_i m_{i,1} - d_i \varepsilon + a_{ii} f(m_{i,2} + \varepsilon) + \sum_{j=1, j \neq i}^n \max\{a_{ij} u_j, a_{ij} v_j\} + J_i < 0,$$

$$F_i(n_{i,2} - \varepsilon) \geq -d_i n_{i,3} + d_i \varepsilon + a_{ii} f_i(n_{i,2} - \varepsilon) + \sum_{j=1, j \neq i}^n \min\{a_{ij} u_j, a_{ij} v_j\} + J_i > 0.$$

由函数 $F_i(\xi)$ 的连续性可知, 存在一点 $\bar{x}_i \in (m_{i,2} + \varepsilon, n_{i,2} - \varepsilon)$, 使得 $F_i(\bar{x}_i) = 0$ 。

情况 5 当 $i \in N_5$ 时, 已知 $f_i(m_{i,3}) = u_i$, 并且 $f_i(n_{i,2}) = f_i(n_{i,3}) = v_i, m_{i,1} < m_{i,3}, n_{i,2} < n_{i,3}$, 由定理 1 中的式(4)和式(5)可以得到:

$$F_i(n_{i,2} + \varepsilon) > -d_i n_{i,3} - d_i \varepsilon + a_{ii} f_i(n_{i,2} + \varepsilon) + \sum_{j=1, j \neq i}^n \min\{a_{ij} u_j, a_{ij} v_j\} + J_i \geq 0,$$

$$F_i(m_{i,3} - \varepsilon) < -d_i m_{i,1} + d_i \varepsilon + a_{ii} f_i(m_{i,3} - \varepsilon) + \sum_{j=1, j \neq i}^n \max\{a_{ij} u_j, a_{ij} v_j\} + J_i \leq 0.$$

由函数 $F_i(\xi)$ 的连续性可知,存在一点 $\bar{x}_i \in [n_{i,2} + \epsilon, m_{i,3} - \epsilon]$,使得 $F_i(\bar{x}_i) = 0$ 。

情况 6 当 $i \in N_6$ 时,已知 $f_i(m_{i,3}) = u_i, m_{i,1} < m_{i,3}$,并且 $f_i(n_{i,3}) = v_i$ 。由定理 1 中的式(4)和式(5),可以得到不等式:

$$F_i(m_{i,3} + \epsilon) < -d_i m_{i,1} - d_i \epsilon + a_{ii} f(m_{i,3} + \epsilon) + \sum_{j=1, j \neq i}^n \max\{a_{ij} u_j, a_{ij} v_j\} + J_i < 0,$$

$$F_i(n_{i,3} - \epsilon) \geq -d_i n_{i,3} + d_i \epsilon + a_{ii} f_i(n_{i,3} - \epsilon) + \sum_{j=1, j \neq i}^n \min\{a_{ij} u_j, a_{ij} v_j\} + J_i > 0.$$

由函数 $F_i(\xi)$ 的连续性可知,存在一点 $\bar{x}_i \in (m_{i,2} + \epsilon, n_{i,2} - \epsilon)$,使得 $F_i(\bar{x}_i) = 0$ 。

情况 7 当 $i \in N_7$ 时,由已知条件可以得到 $v_i = f_i(n_{i,3} + \epsilon)$,并且 $u_i \leq f_j(x_j) \leq v_j$,由 ϵ 的定义(式(3))与定理 1 中的式(5)可以得到:

$$F_i(n_{i,3} + \epsilon) \geq -d_i n_{i,3} - d_i \epsilon + a_{ii} v_i + \sum_{j=1, j \neq i}^n \min\{a_{ij} u_j, a_{ij} v_j\} + J_i > 0, F_i\left(\frac{1}{\epsilon}\right) \leq -\frac{d_i}{\epsilon} + \sum_{j=1}^n |a_{ij}| \mu_j + |J_i| < 0.$$

由函数 $F_i(\xi)$ 的连续性可知,存在一点 $\bar{x}_i \in \left[n_{i,3} + \epsilon, \frac{1}{\epsilon}\right)$,使得 $F_i(\bar{x}_i) = 0$ 。

接下来,定义一个映射 $M: \tilde{\Omega}_\epsilon \rightarrow \tilde{\Omega}_\epsilon$,其中 $M(x_1, x_2, \dots, x_n) = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)$,很明显可以得到映射 M 是连续的。通过 Brouwer 不动点定理,存在一点 $\mathbf{x}^* = (x_1^*, x_2^*, \dots, x_n^*)^T \in \tilde{\Omega}_\epsilon$,且满足 $M(x_1^*, x_2^*, \dots, x_n^*) = (x_1^*, x_2^*, \dots, x_n^*)$,即以下等式成立:

$$-d_i x_i^* + \sum_{j=1}^n a_{ij} f_j(x_j^*) + J_i = 0.$$

综上所述,可以得到 \mathbf{x}^* 为神经网络(式(2))在区域 $\tilde{\Omega}_\epsilon$ 中的一个平衡点,由于区域 $\tilde{\Omega}_\epsilon$ 选取的任意性,在整个区域 Ω_ϵ 中,该神经网络至少存在 7^n 个平衡点。

得到定理 1 后,可以通过定理 1 证明命题 1,该命题所得到的结论将会在后面定理的证明中使用。

命题 1 由定理 1 中的式(4)与式(5)可以得到以下不等式成立:

$$a_{ii} > 0, -d_i + a_{ii} l_{i,1} > 0, -d_i + a_{ii} l_{i,2} < 0, -d_i + a_{ii} l_{i,3} > 0, -d_i + a_{ii} l_{i,4} < 0, -d_i + a_{ii} l_{i,5} > 0.$$

证明:由式(4)与式(5),并且 $f(n_{i,1}) = f(n_{i,2}), n_{i,1} < n_{i,2}$,可以证得以下不等式成立:

$$-d_i m_{i,1} + a_{ii} u_i < -d_i n_{i,2} + a_{ii} f_i(n_{i,2}) < -d_i n_{i,1} + a_{ii} f_i(n_{i,1}),$$

则有 $d_i(n_{i,1} - m_{i,1}) < a_{ii}(f_i(n_{i,1}) - u_i)$ 。由式(2)设定条件,已知 $d_i > 0, n_{i,1} > m_{i,1}, f_i(n_{i,1}) > u_i$,故可以得到 $a_{ii} > 0$;由条件 $u_i = f(m_{i,1}), l_{i,1} = (f_i(n_{i,1}) - f_i(m_{i,1})) / (n_{i,1} - m_{i,1})$,可以证得 $-d_i + a_{ii} l_{i,1} > 0$ 。

同理,由式(4)与式(5),以及条件 $f(m_{i,1}) = f(m_{i,2}), m_{i,1} < m_{i,2}$,可以证得 $-d_i + a_{ii} l_{i,3} > 0$ 。

上述证明已经得到: $a_{ii} > 0, -d_i + a_{ii} l_{i,1} > 0, -d_i + a_{ii} l_{i,3} > 0$,并且已知 $d_i > 0, l_{i,2} < 0$,由上述条件可以证得 $-d_i + a_{ii} l_{i,2} < 0$ 。同理可证得 $-d_i + a_{ii} l_{i,4} < 0, -d_i + a_{ii} l_{i,5} > 0$ 。

2.2 平衡点的稳定性分析

给定一个集合:

$$\Phi = \left\{ \prod_{i=1}^n \left[\frac{-1}{\epsilon}, m_{i,1} - \epsilon \right]^{\delta_1^{(i)}} \times [n_{i,1} + \epsilon, m_{i,2} - \epsilon]^{\delta_2^{(i)}} \times [n_{i,2} + \epsilon, m_{i,3} - \epsilon]^{\delta_3^{(i)}} \times \left[n_{i,3} + \epsilon, \frac{1}{\epsilon} \right]^{\delta_4^{(i)}}, \right. \\ \left. (\delta_1^{(i)}, \delta_2^{(i)}, \delta_3^{(i)}, \delta_4^{(i)}) = (1, 0, 0, 0) \text{ 或 } (0, 1, 0, 0) \text{ 或 } (0, 0, 1, 0) \text{ 或 } (0, 0, 0, 1) \right\}.$$

同时定义集合 $\Psi = \Omega_\epsilon - \Phi$ 。据此,可以明显得到集合 Φ 有 4^n 个区域,则集合 Ψ 有 $7^n - 4^n$ 个区域。

接下来,给出定理 2 证明集合 Φ 为神经网络(式(2))的正向不变集。

定理 2 如果定理 1 中的式(4)和式(5)成立,则集合 Φ 为神经网络(2)的正向不变集。

证明:任意给定一个集合 $\tilde{\Phi} = \prod_{i \in N_1} \left[-\frac{1}{\epsilon}, m_{i,1} - \epsilon \right] \times \prod_{i \in N_2} [n_{i,1} + \epsilon, m_{i,2} - \epsilon] \times \prod_{i \in N_3} [n_{i,2} + \epsilon, m_{i,3} - \epsilon] \times$

$$\prod_{i \in N_4} \left[n_{i,3} + \epsilon, \frac{1}{\epsilon} \right] \subset \Phi.$$

式中: N_i 为集合 $\{1, 2, \dots, n\}$ 的子集, $N_i \cap N_j = \emptyset (i \neq j, i, j = 1, 2, 3, 4)$, 并且 $N_1 \cup N_2 \cup N_3 \cup N_4 = \{i = 1, 2, \dots, n\}$ 。若集合 $\tilde{\Phi}$ 为正向不变集, 那么对任意初值 $x(0) \in \tilde{\Phi}$, 则其相应的解 $x(t; x(0)) \in \tilde{\Phi}$, 其中 $t > 0$ 。

接下来, 采用反证法对定理 2 进行证明。假设集合 Φ 不是正向不变集, 则有以下 4 种情况需要讨论。

情况 1 当 $i \in N_1$ 时, 假设存在 $t^* > 0$, 使得 $x_i(t^*) = m_{i,1} - \epsilon, \dot{x}_i(t^*) > 0$, 且对于 $0 \leq t \leq t^*$, 有 $x_i(t) \in \left[-\frac{1}{\epsilon}, m_{i,1} - \epsilon\right]$, 其中 ϵ 为一个足够小的常数。由定理 1 中的式(4)与 $f_i(x)$ 的定义可以得到:

$$\begin{aligned} \dot{x}_i(t^*) &= -d_i x_i(t^*) + a_{ii} f_i(x_i(t^*)) + \sum_{j=1, j \neq i}^n a_{ij} f_j(x_j(t^*)) + J_i \leq \\ &-d_i m_{i,1} + a_{ii} u_i + d_i \epsilon + \sum_{j=1, j \neq i}^n \max\{a_{ij} u_j, a_{ij} v_j\} + J_i \leq 0. \end{aligned}$$

这与假设中的 $\dot{x}_i(t^*) > 0$ 矛盾, 故当 $i \in N_1$ 时, 定理 2 成立。

情况 2 当 $i \in N_2$ 时, 分为两种情况进行假设。假设存在 $t^* > 0$, 当 $x_i(t^*) = n_{i,1} + \epsilon$ 时, $\dot{x}_i(t^*) < 0$, 对于 $0 \leq t \leq t^*$, 有 $x_i(t) \in [n_{i,1} + \epsilon, m_{i,2} - \epsilon]$; 或是当 $x_i(t^*) = m_{i,2} - \epsilon, \dot{x}_i(t^*) > 0$, 对于 $0 \leq t \leq t^*$, 有 $x_i(t) \in [n_{i,1} + \epsilon, m_{i,2} - \epsilon]$ 。 ϵ 为一个足够小的常数。

首先对第一种假设进行证明。通过命题 1 中的证明可知 $f(n_{i,1} + \epsilon) = f(n_{i,1}) + l_{i,2}\epsilon, d_i - a_{ii}l_{i,2} > 0$, 再由定理 1 中的式(5)与 $f_i(x)$ 的定义可以得到:

$$\begin{aligned} \dot{x}_i(t^*) &= -d_i x_i(t^*) + a_{ii} f_i(x_i(t^*)) + \sum_{j=1, j \neq i}^n a_{ij} f_j(x_j(t^*)) + J_i \geq \\ &-d_i n_{i,1} + a_{ii} v_i - (d_i - a_{ii}l_{i,2})\epsilon + \sum_{j=1, j \neq i}^n \min\{a_{ij} u_j, a_{ij} v_j\} + J_i \geq 0. \end{aligned}$$

这与假设中的 $\dot{x}_i(t^*) < 0$ 矛盾。第二种假设可以类似地证明, 故当 $i \in N_2$ 时, 定理 2 成立。

情况 3 当 $i \in N_3$ 时, 同样分为两种情况进行假设。假设存在 $t^* > 0$, 当 $x_i(t^*) = n_{i,2} + \epsilon$ 时, $\dot{x}_i(t^*) < 0$, 对于 $0 \leq t \leq t^*$, 有 $x_i(t) \in [n_{i,2} + \epsilon, m_{i,3} - \epsilon]$; 或是当 $x_i(t^*) = m_{i,3} - \epsilon$ 时, $\dot{x}_i(t^*) > 0$, 对于 $0 \leq t \leq t^*$, 有 $x_i(t) \in [n_{i,2} + \epsilon, m_{i,3} - \epsilon]$ 。其中 ϵ 为一个足够小的常数。

首先对第一种假设进行证明。通过命题 1 中的结论可知: $f(n_{i,2} + \epsilon) = f(n_{i,2}) + l_{i,4}\epsilon, d_i - a_{ii}l_{i,4} > 0$, 再由定理 1 中的式(5)与 $f_i(x)$ 的定义可以得到:

$$\begin{aligned} \dot{x}_i(t^*) &= -d_i x_i(t^*) + a_{ii} f_i(x_i(t^*)) + \sum_{j=1, j \neq i}^n a_{ij} f_j(x_j(t^*)) + J_i \geq \\ &-d_i n_{i,2} + a_{ii} v_i - (d_i - a_{ii}l_{i,4})\epsilon + \sum_{j=1, j \neq i}^n \min\{a_{ij} u_j, a_{ij} v_j\} + J_i \geq 0. \end{aligned}$$

这与假设中 $\dot{x}_i(t^*) < 0$ 矛盾。第二种假设同理可证, 故当 $i \in N_3$ 时, 定理 2 成立。

情况 4 当 $i \in N_4$ 时, 假设存在 $t^* > 0$ 使得 $x_i(t^*) = n_{i,3} + \epsilon, \dot{x}_i(t^*) < 0$, 并且对于 $0 \leq t \leq t^*$, 有 $x_i(t) \in \left[n_{i,3} + \epsilon, \frac{1}{\epsilon}\right]$ 。由定理 1 中的式(5)与 $f_i(x)$ 的定义可以得到:

$$\begin{aligned} \dot{x}_i(t^*) &= -d_i x_i(t^*) + a_{ii} f_i(x_i(t^*)) + \sum_{j=1, j \neq i}^n a_{ij} f_j(x_j(t^*)) + J_i \geq \\ &-d_i n_{i,3} + a_{ii} v_i - d_i \epsilon + \sum_{j=1, j \neq i}^n \min\{a_{ij} u_j, a_{ij} v_j\} + J_i \geq 0. \end{aligned}$$

上述结果与假设中的 $\dot{x}_i(t^*) < 0$ 矛盾, 故当 $i \in N_4$ 时, 定理 2 成立。

综上所述, 由于子集 $\tilde{\Phi}$ 选取的任意性, 可以证明集合 Φ 为神经网络(式(2))的正向不变集。

给出神经网络(式(2))的正向不变集后, 接下来通过定理 3 和定理 4 对神经网络平衡点的多稳定性以及不稳定性进行分析。

定理 3 如果基于连续锯齿型激活函数(式(1))的递归神经网络(式(2))同时满足定理 1 中的式(4)和式(5)以及条件

$$\sum_{j=1, j \neq i}^n |a_{ij}| \max\{|l_{j,1}|, |l_{j,2}|, |l_{j,3}|, |l_{j,4}|, |l_{j,5}|\} < \min\{d_i, |a_{ii}l_{i,1} - d_i|, |a_{ii}l_{i,2} - d_i|, |a_{ii}l_{i,3} - d_i|, |a_{ii}l_{i,4} - d_i|, |a_{ii}l_{i,5} - d_i|\}, \quad (6)$$

那么神经网络(式(2))至少有 7^n 个平衡点,其中 4^n 个平衡点是局部指数稳定的,其余 $7^n - 4^n$ 个平衡点是不稳定的。接下来对每一个平衡点的稳定性进行讨论。

情况 1 任意给定一个子集

$$\tilde{\Phi} = \prod_{i \in N_1} \left[\frac{-1}{\epsilon}, m_{i,1} - \epsilon \right] \times \prod_{i \in N_2} [n_{i,1} + \epsilon, m_{i,2} - \epsilon] \times \prod_{i \in N_3} [n_{i,2} + \epsilon, m_{i,3} - \epsilon] \times \prod_{i \in N_4} \left[n_{i,3} + \epsilon, \frac{1}{\epsilon} \right] \subset \Phi.$$

式中: N_i 为集合 $\{1, 2, \dots, n\}$ 的子集, $N_i \cap N_j = \emptyset (i \neq j, i, j = 1, 2, 3, 4)$, $N_1 \cup N_2 \cup N_3 \cup N_4 = \{1, 2, \dots, n\}$,集合 Φ 的定义同定理 2。

将 $\mathbf{x}(t; \mathbf{x}(0))$ 定义为初值条件为 $\mathbf{x}(0) \in \tilde{\Phi}$ 的式(2)的解。再设 \mathbf{x}^* 为区域 $\tilde{\Phi}$ 的一个平衡点。作变换 $\boldsymbol{\gamma}(t) = \mathbf{x}(t; \mathbf{x}(0)) - \mathbf{x}^*$, 由定理 2 可知 $\tilde{\Phi}$ 为一个正向不变集,即对任意 $\mathbf{x}(0) \in \tilde{\Phi}$, 式(2)的解 $\mathbf{x}(t; \mathbf{x}(0)) \in \tilde{\Phi}$ 。当 $\mathbf{x}(t; \mathbf{x}(0)) \in \Phi$ 时, $\sum_{j \in N_1} f_j(x_j)$ 与 $\sum_{j \in N_4} f_j(x_j)$ 均为常数,因此 $\sum_{j \in N_1} f_j(x_j) = \sum_{j \in N_1} f_j(x_j^*)$ 与 $\sum_{j \in N_4} f_j(x_j) = \sum_{j \in N_4} f_j(x_j^*)$ 成立,可得:

$$\dot{\gamma}_i(t) = -d_i \gamma_i(t) + \sum_{j \in N_2} a_{ij} [f_j(x_j^* + \gamma_j(t)) - f_j(x_j^*)] + \sum_{j \in N_3} a_{ij} [f_j(x_j^* + \gamma_j(t)) - f_j(x_j^*)]. \quad (7)$$

不失一般性,假设 $N_2 = \{1, 2, \dots, p\}$, $N_3 = \{p+1, p+2, \dots, q\}$, $p \leq q \leq n$, 可将式(7)简化为:

$$\dot{\gamma}_i(t) = -d_i \gamma_i(t) + \sum_{j=1}^p a_{ij} l_{j,2} \gamma_j(t) + \sum_{j=p+1}^q a_{ij} l_{j,4} \gamma_j(t). \quad (8)$$

因此,可以得到式(8)的系数矩阵

$$\hat{\mathbf{A}}_1 = \begin{pmatrix} -d_1 + a_{11}l_{1,2} & a_{12}l_{2,2} & \cdots & a_{1p}l_{p,2} & a_{1(p+1)}l_{(p+1),4} & \cdots & a_{1q}l_{q,4} & 0 & \cdots & 0 \\ a_{21}l_{1,2} & -d_2 + a_{22}l_{2,2} & \cdots & a_{2p}l_{p,2} & a_{2(p+1)}l_{(p+1),4} & \cdots & a_{2q}l_{q,4} & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{p1}l_{1,2} & a_{p2}l_{2,2} & \cdots & -d_p + a_{pp}l_{p,2} & a_{p(p+1)}l_{(p+1),4} & \cdots & a_{pq}l_{q,4} & 0 & \cdots & 0 \\ a_{(p+1)1}l_{1,2} & a_{(p+1)2}l_{2,2} & \cdots & a_{(p+1)p}l_{p,2} & -d_{p+1} + a_{(p+1)(p+1)}l_{(p+1),4} & \cdots & a_{(p+1)q}l_{q,4} & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{q1}l_{1,2} & a_{q2}l_{2,2} & \cdots & a_{qp}l_{p,2} & a_{q(p+1)}l_{(p+1),4} & \cdots & -d_q + a_{qq}l_{q,4} & 0 & \cdots & 0 \\ a_{(q+1)1}l_{1,2} & a_{(q+1)2}l_{2,2} & \cdots & a_{(q+1)p}l_{p,2} & a_{(q+1)(p+1)}l_{(p+1),4} & \cdots & a_{(q+1)q}l_{q,4} & -d_{q+1} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1}l_{1,2} & a_{n2}l_{2,2} & \cdots & a_{np}l_{p,2} & a_{n(p+1)}l_{(p+1),4} & \cdots & a_{nq}l_{q,4} & 0 & \cdots & -d_n \end{pmatrix}.$$

通过式(6)以及定义 2,可知 $\hat{\mathbf{A}}_1$ 为对角占优矩阵,再由对角占优矩阵理论以及命题 1 可知 $\hat{\mathbf{A}}_1$ 所有特征值都具有负实部。因此,由局部稳定性判定定理可得到平衡点 \mathbf{x}^* 是 $\tilde{\Phi}$ 的局部指数稳定点。由于对 $\tilde{\Phi}$ 选取的任意性, Φ 的每个子集中都存在一个稳定的平衡点,即神经网络(式(2))在 Φ 中存在 4^n 个局部指数稳定的平衡点。

情况 2 任意给定一个子集

$$\tilde{\Psi} = \prod_{i \in N_1} \left[\frac{-1}{\epsilon}, m_{i,1} - \epsilon \right] \times \prod_{i \in N_2} [m_{i,1} + \epsilon, n_{i,1} - \epsilon] \times \prod_{i \in N_3} [n_{i,1} + \epsilon, m_{i,2} - \epsilon] \times \prod_{i \in N_4} [m_{i,2} + \epsilon, n_{i,2} - \epsilon] \times \prod_{i \in N_5} [n_{i,2} + \epsilon, m_{i,3} - \epsilon] \times \prod_{i \in N_6} [m_{i,3} + \epsilon, n_{i,3} - \epsilon] \times \prod_{i \in N_7} \left[n_{i,3} + \epsilon, \frac{1}{\epsilon} \right] \subset \Psi.$$

式中: $\Psi = \Omega_\epsilon - \Phi$, 集合 Φ 的定义同定理 2, N_i 为集合 $\{1, 2, \dots, n\}$ 的子集, $N_1 \cup N_2 \cup N_3 \cup N_4 \cup N_5 \cup N_6 \cup N_7 = \{1, 2, \dots, n\}$, 并且 $N_i \cap N_j = \emptyset (i \neq j, i, j = 1, 2, \dots, 7)$, $N_2 \cup N_4 \neq \emptyset$ 。与情况 1 的证明类似, 设 $\mathbf{x}(t; \mathbf{x}(0))$ 为式(2)中点 \mathbf{x}^* 附近的一个解, 作变换 $\boldsymbol{\gamma}(t) = \mathbf{x}(t; \mathbf{x}(0)) - \mathbf{x}^*$, 可以得到等式:

$$\begin{aligned} \dot{\gamma}_i(t) = & -d_i \gamma_i(t) + \sum_{j \in N_2} a_{ij} [f_j(x_j^* + \gamma_j(t)) - f_j(x_j^*)] + \sum_{j \in N_4} a_{ij} [f_j(x_j^* + \gamma_j(t)) - f_j(x_j^*)] + \\ & \sum_{j \in N_6} a_{ij} [f_j(x_j^* + \gamma_j(t)) - f_j(x_j^*)] + \sum_{j \in N_3} a_{ij} [f_j(x_j^* + \gamma_j(t)) - f_j(x_j^*)] + \\ & \sum_{j \in N_5} a_{ij} [f_j(x_j^* + \gamma_j(t)) - f_j(x_j^*)]。 \end{aligned} \quad (9)$$

记 $N_2 = \{1, 2, \dots, p\}$, $N_4 = \{p+1, p+2, \dots, q\}$, $N_6 = \{q+1, q+2, \dots, r\}$, $N_3 = \{r+1, r+2, \dots, s\}$, $N_5 = \{s+1, s+2, \dots, w\}$, 则可以将式(9)化简为:

$$\begin{aligned} \dot{\gamma}_i(t) = & -d_i \gamma_i(t) + \sum_{j=1}^p a_{ij} l_{j,1} \gamma_j(t) + \sum_{j=p+1}^q a_{ij} l_{j,3} \gamma_j(t) + \sum_{j=q+1}^r a_{ij} l_{j,5} \gamma_j(t) + \\ & \sum_{j=r+1}^s a_{ij} l_{j,2} \gamma_j(t) + \sum_{j=s+1}^w a_{ij} l_{j,4} \gamma_j(t)。 \end{aligned} \quad (10)$$

可得到式(9)的系数矩阵

$$\widehat{\mathbf{A}}_2 = \begin{pmatrix} -d_1 + a_{11} l_{1,1} & a_{12} l_{2,1} & \dots & a_{1p} l_{p,1} & \dots & a_{1q} l_{q,3} & \dots & a_{1r} l_{r,5} & \dots & a_{1s} l_{s,2} & \dots & a_{1w} l_{w,4} & 0 & \dots & 0 \\ a_{21} l_{1,1} & -d_2 + a_{22} l_{2,1} & \dots & a_{2p} l_{p,1} & \dots & a_{2q} l_{q,3} & \dots & a_{2r} l_{r,5} & \dots & a_{2s} l_{s,2} & \dots & a_{2w} l_{w,4} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{p1} l_{1,1} & a_{p2} l_{2,1} & \dots & -d_p + a_{pp} l_{p,1} & \dots & a_{pq} l_{q,3} & \dots & a_{pr} l_{r,5} & \dots & a_{ps} l_{s,2} & \dots & a_{pw} l_{w,4} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{q1} l_{1,1} & a_{q2} l_{2,1} & \dots & a_{qp} l_{p,1} & \dots & -d_q + a_{qq} l_{q,3} & \dots & a_{qr} l_{r,5} & \dots & a_{qs} l_{s,2} & \dots & a_{qw} l_{w,4} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{r1} l_{1,1} & a_{r2} l_{2,1} & \dots & a_{rp} l_{p,1} & \dots & a_{rq} l_{q,3} & \dots & -d_r + a_{rr} l_{r,5} & \dots & a_{rs} l_{s,2} & \dots & a_{rw} l_{w,4} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{s1} l_{1,1} & a_{s2} l_{2,1} & \dots & a_{sp} l_{p,1} & \dots & a_{sq} l_{q,3} & \dots & a_{sr} l_{r,5} & \dots & -d_s + a_{ss} l_{s,2} & \dots & a_{sw} l_{w,4} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{w1} l_{1,1} & a_{w2} l_{2,1} & \dots & a_{wp} l_{p,1} & \dots & a_{wq} l_{q,3} & \dots & a_{wr} l_{r,5} & \dots & a_{ws} l_{s,2} & \dots & -d_w + a_{ww} l_{w,4} & 0 & \dots & 0 \\ a_{(w+1)1} l_{1,1} & a_{(w+1)2} l_{2,1} & \dots & a_{(w+1)p} l_{p,1} & \dots & a_{(w+1)q} l_{q,3} & \dots & a_{(w+1)r} l_{r,5} & \dots & a_{(w+1)s} l_{s,2} & \dots & a_{(w+1)w} l_{w,4} & -d_{w+1} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} l_{1,1} & a_{n2} l_{2,1} & \dots & a_{np} l_{p,1} & \dots & a_{nq} l_{q,3} & \dots & a_{nr} l_{r,5} & \dots & a_{ns} l_{s,2} & \dots & a_{nw} l_{w,4} & 0 & \dots & -d_n \end{pmatrix}。$$

通过式(6)及定义 2 可知 $\widehat{\mathbf{A}}_2$ 为对角占优矩阵。再由对角占优矩阵理论及命题 1 可知 $\widehat{\mathbf{A}}_2$ 具有 r 个正实部特征值, 其余 $n-r$ 个特征值均为负实部特征值, 因此 \mathbf{x}^* 为 $\widehat{\Psi}$ 中的一个不稳定平衡点。由于 $\widehat{\Psi}$ 选取的任意性, 在 Ψ 中的每一个子集都存在一个不稳定的平衡点, 即神经网络(式(2))具有 $7^n - 4^n$ 个不稳定的平衡点。

综上所述, 具有连续锯齿型激活函数(式(1))的递归神经网络(式(2))在满足式(4)~式(6)的情况下, 至少具有 7^n 个平衡点, 其中 4^n 个平衡点局部指数稳定, 剩余的 $7^n - 4^n$ 个平衡点不稳定。定理 3 证毕。

2.3 激活函数的推广

给出基于连续锯齿型激活函数(式(1))的递归神经网络(式(2))的多稳定性分析的全部证明后, 本研究对证明结果进行推广。分析过程中发现, 连续锯齿型激活函数峰值点的数目与局部稳定平衡点的数量之间存在定量关系, 因此本研究增加了激活函数峰值点的数目, 提出推广后的激活函数:

$$f_i(x) = \begin{cases} u_i, & -\infty < x < m_{i,1}; \\ l_{i,1}x + c_{i,1}, & m_{i,1} \leq x < n_{i,1}; \\ l_{i,2}x + c_{i,2}, & n_{i,1} \leq x < m_{i,2}; \\ \dots & \dots \\ l_{i,2k-3}x + c_{i,2k-3}, & m_{i,k-1} \leq x < n_{i,k-1}; \\ l_{i,2k-2}x + c_{i,2k-2}, & n_{i,k-1} \leq x < m_{i,k}; \\ l_{i,2k-1}x + c_{i,2k-1}, & m_{i,k} \leq x < n_{i,k}; \\ v_i, & n_{i,k} \leq x < +\infty. \end{cases} \quad (11)$$

式中: $m_{i,1}, m_{i,2}, \dots, m_{i,k}, n_{i,1}, n_{i,2}, \dots, n_{i,k}, u_i, v_i$ 均为常数,且满足 $-\infty < m_{i,1} < n_{i,1} < m_{i,2} < \dots < n_{i,k} < m_{i,k} < n_{i,k} < +\infty$, $u_i = f_i(m_{i,1}) = f_i(m_{i,2}) = \dots = f_i(m_{i,k})$, $v_i = f_i(n_{i,1}) = f_i(n_{i,2}) = \dots = f_i(n_{i,k})$, 且 $u_i < v_i, i = 1, 2, \dots, n$; $l_{i,1}, l_{i,2}, \dots, l_{i,2k-1}, c_{i,1}, c_{i,2}, \dots, c_{i,2k-1}$ 也为常数,并且 $l_{i,p}$ 满足条件:若 p 为奇数,则 $l_{i,p} > 0$;若 p 为偶数,则 $l_{i,p} < 0$,其中 $1 \leq p \leq 2k - 1, i = 1, 2, \dots, n$ 。在推广的连续锯齿型激活函数(式(11))中存在 k 个峰值点,分别为: $(n_{i,1}, f_i(n_{i,1})), (n_{i,2}, f_i(n_{i,2})), \dots, (n_{i,k}, f_i(n_{i,k}))$ 。

基于激活函数(式(11))提出以下定理。

定理 4 如果对任意 $i = 1, 2, \dots, n$, 若均满足条件

$$-d_i m_{i,1} + a_{ii} u_i + \sum_{j=1, j \neq i}^n \max\{a_{ij} u_j, a_{ij} v_j\} + J_i < 0, \quad (12)$$

$$-d_i n_{i,k} + a_{ii} v_i + \sum_{j=1, j \neq i}^n \min\{a_{ij} u_j, a_{ij} v_j\} + J_i > 0, \quad (13)$$

则基于激活函数(式(11))的神经网络(式(2))至少具有 $(2k + 1)^n$ 个平衡点。

定理 4 的证明过程与定理 1 的证明过程类似,此处证明略。

定理 5 如果基于连续锯齿型激活函数(式(1))的递归神经网络(式(2))同时满足定理 4 中的式(12)和式(13)以及条件

$$\sum_{j=1, j \neq i}^n |a_{ij}| \max\{|l_{j,r}|\} < \min\{d_i, |a_{ii} l_{i,r} - d_i|\}, r = 1, 2, \dots, 2k - 1. \quad (14)$$

那么神经网络(式(2))共有 $(2k + 1)^n$ 个平衡点,其中 $(k + 1)^n$ 个平衡点是局部指数稳定的,其余 $(2k + 1)^n - (k + 1)^n$ 个平衡点是不稳定的。

定理 5 的证明过程与定理 3 的证明过程类似,此处证明略。

与 Sigmoid 型函数^[8]、饱和函数^[9]、墨西哥帽型函数^[10]以及高斯小波型激活函数^[11]相比,本研究设计的连续锯齿型激活函数的 Hopfield 神经网络可以产生更多的稳定平衡点,这意味着神经网络的储存容量得到了提升,在联想记忆的应用中可以储存更多的记忆模式。

通过控制激活函数中的参数,使得激活函数的参数满足 $v_i = f(n_{i,1}) = f(n_{i,2}) = \dots = f(n_{i,k})$,与文献[11]相比,本研究满足 n 维神经网络多平衡点存在性的条件数由 $3n$ 减少到 $2n$,而且随着激活函数中峰值点的增加,满足多平衡点存在性的条件数依然为 $2n$,从而大大降低了计算复杂度。

本研究的创新点在于给出了一个连续锯齿型激活函数(式(1)),讨论了基于该激活函数的递归神经网络(式(2))的多平衡点的存在性与稳定性,并给出了定理 1 与定理 3。研究发现,连续锯齿型激活函数峰值点与神经网络(式(2))的稳定平衡点数之间存在定量关系,于是将激活函数(式(1))推广为具有 k 个峰值的连续锯齿型激活函数(式(11))。由定理 4 与定理 5 给出的结果可以看出,本研究提出的通过增加激活函数峰值点的数目来增加神经网络稳定平衡点的方法,操作灵活且方便快捷。在实际应用中,将神经网络应用到联想记忆或模式识别时,可以通过选取具有适当峰值点的连续锯齿型激活函数的神经网络来获取具有合适存储容量的神经网络。

3 数值算例

例 1 考虑具有两个神经元的神经网络

$$\begin{cases} \dot{x}_1(t) = -x_1(t) + 8.5f(x_1(t)) + 0.2f(x_2(t)) + 1.5, \\ \dot{x}_2(t) = -x_2(t) - 0.2f(x_1(t)) + 8.5f(x_2(t)) + 1.5. \end{cases} \quad (15)$$

激活函数定义为:

$$f(x) = \begin{cases} -1, & -\infty < x < -1; \\ x, & -1 \leq x < 1; \\ -x + 2, & 1 \leq x < 3; \\ x - 4, & 3 \leq x < 5; \\ -x + 6, & 5 \leq x < 7; \\ x - 8, & 7 \leq x < 9; \\ 1, & 9 \leq x < +\infty. \end{cases} \quad (16)$$

如图 1 所示,激活函数(式(16))具有 3 个峰值点 (1,1)、(5,1)、(9,1)。根据定理 1 与定理 3 提出的充分条件可知,基于激活函数(式(16))的神经网络模型(式(15))满足定理 3 给出的条件。因此,神经网络(式(15))至少具有 $7^2 = 49$ 个平衡点,其中 $4^2 = 16$ 个平衡点是局部稳定的,其余 $7^2 - 4^2 = 33$ 个平衡点为不稳定平衡点。随机给出 500 个初始条件 $\mathbf{x}(0) = (x_1(0), x_2(0))^T$, 初始条件从区域 $[-9, 11] \times [-9, 11]$ 选取,可以得到系统(式(15))的相图如图 2 所示,很明显图 2 中存在 16 个稳定平衡点。

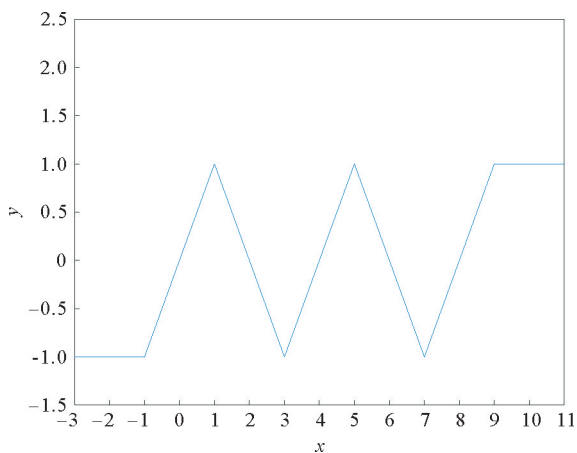


图 1 例 1 中激活函数(式(16))的图像

Fig. 1 Image of activation function (formula(16)) in Example 1

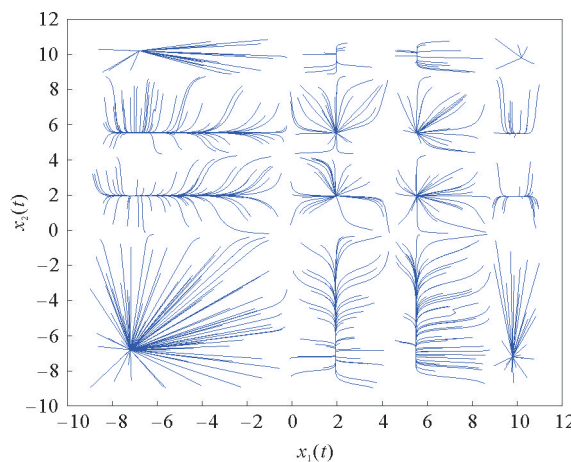


图 2 例 1 中 500 个随机初始条件下系统(式(15))的相图

Fig. 2 Phase diagram of system (formula(15)) with 500 random initial conditions in Example 1

例 2 考虑具有两个神经元的神经网络

$$\begin{cases} \dot{x}_1(t) = -x_1(t) + 11.5f(x_1(t)) + 0.2f(x_2(t)) + 2.5, \\ \dot{x}_2(t) = -x_2(t) - 0.2f(x_1(t)) + 11.5f(x_2(t)) + 2.5. \end{cases} \quad (17)$$

激活函数定义为:

$$f(x) = \begin{cases} -1, & -\infty < x < -1; \\ x, & -1 \leq x < 1; \\ -x + 2, & 1 \leq x < 3; \\ x - 4, & 3 \leq x < 5; \\ -x + 6, & 5 \leq x < 7; \\ x - 8, & 7 \leq x < 9; \\ -x + 10, & 9 \leq x < 11; \\ x - 12, & 11 \leq x < 13; \\ 1, & 13 \leq x < +\infty. \end{cases} \quad (18)$$

如图 3 所示,激活函数(式(18))具有 4 个峰值点 (1,1)、(5,1)、(9,1)、(13,1)。根据定理 4 与定理 5 提出的充分条件,基于激活函数(式(18))的神经网络模型(式(17))满足定理 5 的条件。因此,神经网络(式(17))至少具有 $9^2 = 81$ 个平衡点,其中 $5^2 = 25$ 个平衡点为局部稳定点,其余 $9^2 - 5^2 = 56$ 个平衡点为不稳定平衡点。随机给出 500 个初始条件 $\mathbf{x}(0) = (x_1(0), x_2(0))^T$, 初始条件从区域 $[-9, 16] \times [-9, 16]$ 选取神经网络(式(17))的相图如图 4 所示,可明显看出存在 25 个稳定平衡点。

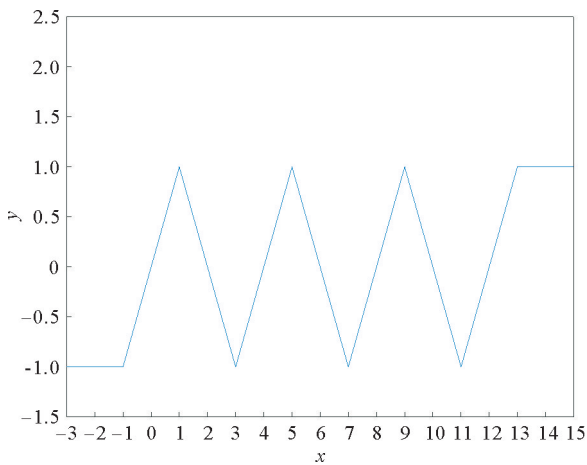


图 3 例 2 中激活函数(式(18))的图像

Fig. 3 Image of activation function (formula(18)) in Example 2

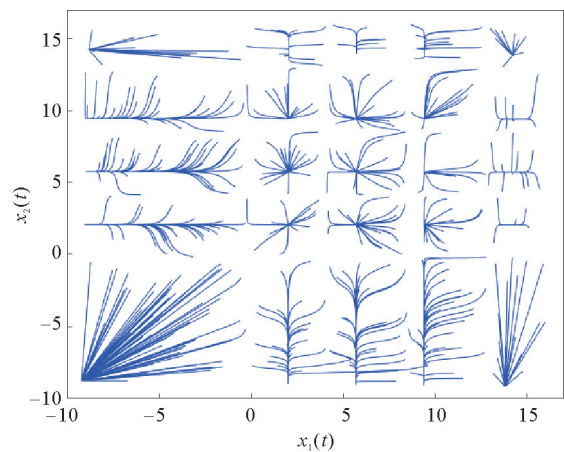


图 4 例 2 中 500 个随机初始条件下系统(式(17))的相图

Fig. 4 Phase diagram of system (formula(17)) with 500 random initial conditions in Example 2

4 结论

本研究设计了一类连续锯齿型激活函数,通过区间剖分法、Brouwer 不动点定理、对角占优矩阵理论等研究了基于该激活函数的 Hopfield 递归神经网络的多稳定性问题,证明了基于连续锯齿型激活函数的 n 维神经网络模型至少具有 7^n 个平衡点,其中 4^n 个平衡点是局部指数稳定的,其余的平衡点是不稳定的。与现有的一些结论相比,该激活函数能产生更多的稳定平衡点。此外,当增加峰值点时,可以得到 n 维神经网络在含有 k 个峰值点的连续锯齿型激活函数中至少具有 $(2k + 1)^n$ 个平衡点,其中 $(k + 1)^n$ 个平衡点为局部稳定点。最后通过两个具体的数值算例验证了本研究结果的有效性。

本研究考虑的是较为理想状态下的不含时滞的 Hopfield 神经网络,但在实际应用中时滞问题是不可避免的,未来将对含有时滞的神经网络模型进一步分析。此外,Cohen-Grossberg 神经网络与竞争型神经网络比传统的 Hopfield 神经网络具有更为复杂的动力学行为,未来将对基于连续锯齿型激活函数的 Cohen-Grossberg 神经网络或竞争型神经网络的多稳定性进行学习和研究。

参考文献:

- [1] DENG K,ZHU S,BAO G,et al. Multistability of dynamic memristor delayed cellular neural networks with application to associative memories [J]. IEEE Transactions on Neural Networks and Learning Systems,2023,34(2):690-702.
- [2] DE MENESES F G A,TELES A S,NUNES M,et al. Neural networks to recognize patterns in topographic images of cortical electrical activity of patients with neurological diseases [J]. Brain Topography,2022,35(4):464-480.
- [3] 姚篮,秦婧,王震. 基于采样控制的时滞神经网络的指数同步[J]. 山东科技大学学报(自然科学版),2020,39(1):76-84.
YAO Lan,QIN Jing,WANG Zhen. Exponential synchronization of delayed neural networks via sampled-data control[J]. Journal of Shandong University of Science and Technology(Natural Science),2020,39(1):76-84.
- [4] NIE X B,LIANG J L,CAO J D. Multistability analysis of competitive neural networks with Gaussian-wavelet-type activation functions and unbounded time-varying delays[J]. Applied Mathematics and Computation,2019,356:449-468.
- [5] POPA C A. Global exponential stability of neutral-type octonion-valued neural networks with time-varying delays[J]. Neurocomputing,2018,309:117-133.
- [6] DUAN L,SHI M,WANG Z Y,et al. Global exponential synchronization of delayed complex-valued recurrent neural networks with discontinuous activations[J]. Neural Processing Letters,2019,50(3):2183-2200.
- [7] XU C J,ZHANG W,AOUITI C,et al. Further analysis on dynamical properties of fractional-order bi-directional associative memory neural networks involving double delays[J]. Mathematical Methods in the Applied Sciences,2022,45(17):11736-11754.
- [8] CHENG C Y,LIN K H,SHIH C W. Multistability in recurrent neural networks[J]. SIAM Journal on Applied Mathematics,2006,66(4):1301-1320.
- [9] ZENG Z,WANG J,LIAO X X. Stability analysis of delayed cellular neural networks described using cloning templates[J]. IEEE Transactions on Circuits and Systems I:Regular Papers,2004,51(11):2313-2324.
- [10] NIE X,ZHENG W X. Multistability and instability of neural networks with discontinuous nonmonotonic piecewise linear activation functions [J]. IEEE Transactions on Neural Networks and Learning Systems,2015,26(11):2901-2913.
- [11] NIE X,CAO J D,FEI S M. Multistability and instability of competitive neural networks with non-monotonic piecewise linear activation functions[J]. Nonlinear Analysis:Real World Applications,2019,45:799-821.

(责任编辑:齐敏华)