时滞可交换四元数神经网络稳定性分析

林东源 [1]，陈晓丰 [1]，孙文涛 [2]，李伟凯 [3]，夏砚楠 [1]
（1. 重庆交通大学数学与统计学院，重庆 400074；2. 西藏大学理学院，西藏 拉萨 850000；
3. 南京航空航天大学计算机科学与技术学院，江苏 南京 210016）

摘　要：针对传统的四元数乘法运算不满足交换律的问题，介绍了一种可交换四元数，同时建立一种交换四元数神经网络（CQVNN），并对 CQVNN 的渐近稳定性进行研究。研究过程中，将 CQVNN 分解成 4 个实值神经网络，通过拓扑度理论、传统的李雅普诺夫稳定性理论以及不等式理论，得到 CQVNN 平衡点的存在唯一性条件以及其平衡点渐近稳定的充分条件。最后，给出一个数值案例，以验证所得结论的有效性。

关键词：交换四元数；时滞神经网络；渐近稳定性
中图分类号：O175.13
文献标识码：A
doi: 10.11959/j.issn.2096−6652.202009

Stability analysis of commutative quaternion valued neural network with time varying delays

LIN Dongyuan[1], CHEN Xiaofeng[1], SUN Wentao[2], LI Weikai[3], XIA Yannan[1]
1. College of Mathematics and Statistics, Chongqing Jiaotong University, Chongqing 400074, China
2. School of Science, Tibet University, Lhasa 850000, China
3. College of Computer Science & Technology, Nanjing University of Aeronautics & Astronautics, Nanjing 210016, China

Abstract: Because the traditional quaternion multiplication does not meet the commutative law, a kind of commutative quaternion was introduced, and a commutative quaternion valued neural network (CQVNN) was established, and the asymptotic stability of the CQVNN was studied. In the course of research, the CQVNN were decomposed into four real-valued neural networks. Then, through the topological degree theory, the traditional Lyapunov stability theory and the inequality theory, the conditions for the existence and uniqueness of the equilibrium point of CQVNN and the linear matrix inequality (LMI) conditions for the asymptotically stable of equilibrium point were obtained. Finally, a numerical example was given to verify the validity of the conclusion.

Key words: commutative quaternion, delayed neural network, asymptotically stability

1　引言

自从生物物理学家 Hopfield[1]在 20 世纪 80 年代首次提出人工神经网络以来，人工神经网络得到了广泛的研究。如今，人工神经网络在联想记忆、图像处理、优化组合等方面得到了广泛的应用[2-5]。在这些领域对人工神经网络系统的稳定性进行研究是十分必要的[6-8]。

近年来，随着数域的扩张，人工神经网络主要有实数神经网络、复数神经网络以及四元数神经网络。四元数由一个实部和 3 个虚部构成，四元数对数据信息的存储能力比实数和复数都要强。一个四

收稿日期：2020−01−14；**修回日期：**2020−02−19
通信作者：孙文涛，550668951@qq.com
基金项目：重庆市自然科学基金资助项目（No.cstc2018jcyjAX0606，No.cstc2017jcyjAX0082）；重庆市教育委员会科学技术研究基金资助项目（No.KJQN201900701）
Foundation Items: Natural Science Foundation of Chongqing (No.cstc2018jcyjAX0606, No.cstc2017jcyjAX0082), Scientific and Technology Research Program of Chongqing Municipal Education Commission (No.KJQN201900701)

元数的数据存储能力是一个实数的 4 倍，是一个复数的 2 倍。因此，四元数神经网络在处理高维数据时比实数神经网络以及复数神经网络都更有优势。例如，在与图像有关的应用中，一种颜色往往由红色（R）、绿色（G）和蓝色（B）3 种基本颜色按照各自的像素值合成，将 R、G、B 值分别对应于纯四元数的 3 个虚部部分 i、j 和 k 的系数时，一种颜色便可以用一个纯虚数唯一表示，一张图片可以用一个纯四元数矩阵唯一表示[9-10]。

然而，之前研究的四元数神经网络涉及的四元数是定义在哈密顿（Hamilton）乘法规则之下的，此规则不满足乘法交换律。为了解决这个问题，本文给出了满足乘法交换律的四元数（以下简称交换四元数），它保留了四元数相对于实数及复数的基本优势，即具有更好的数据存储能力。但在新的乘法规则下，对交换四元数神经网络（commutative quaternion valued neural network，CQVNN）的研究也面临着许多挑战，例如：交换四元数同时存在 3 个共轭，交换四元数的模表示十分复杂，交换四元数不满足三角不等式等。这些局限使得那些适用于实数、复数以及不可交换四元数的不等式理论不适用于交换四元数。为了解决这些问题，本文通过交换四元数的乘法规则将 CQVNN 转化为实数神经网络，进而研究其稳定性。

本文的结构组织如下：第 2 节介绍交换四元数以及交换四元数神经网络，给出一些引理及定义；第 3 节给出交换四元数神经网络渐近稳定的充分条件；第 4 节用一个数值案例验证了得到的结论的有效性；第 5 节对本文进行总结。

2　准备工作

2.1　交换四元数

一个交换四元数可以表示为：

$$q = q^R + q^I i + q^J j + q^K k \in \mathbb{Q} \tag{1}$$

其中，q^R、q^I、q^J、$q^K \in \mathbb{R}$ 分别表示交换四元数的第一部分、第二部分、第三部分以及第四部分的数值，虚单位 i、j、k 满足如下运算规则。

$$\begin{cases} i^2 = k^2 = -1, j^2 = 1 \\ ij = ji = k \\ jk = kj = i \\ ki = ik = -j \\ ijk = -1 \end{cases} \tag{2}$$

交换四元数 q 存在如下 3 种共轭：

$$q^{(1)} = q^R - q^I + q^J - q^K$$
$$q^{(2)} = q^R + q^I - q^J + q^K$$
$$q^{(3)} = q^R - q^I - q^J + q^K$$

q 的模定义如下：

$$\begin{aligned} \|q\|^4 = qq^{(1)}q^{(2)}q^{(3)} = \\ [(q^R + q^J)^2 + (q^I + q^K)^2] \times \\ [(q^R - q^J)^2 + (q^I - q^K)^2] \geqslant 0 \end{aligned}$$

对于另一个交换四元数 $p = p^R + p^I i + p^J j + p^K k$ 来说，p 与 q 的乘积定义如下：

$$\begin{aligned} pq = qp = \\ (p^R q^R - p^I q^I + p^J q^J - p^K q^K) + \\ (p^R q^I + p^I q^R + p^J q^K + p^K q^J) i + \\ (p^R q^J + p^J q^R - p^I q^K - p^K q^I) j + \\ (p^R q^K + p^K q^R + p^I q^J + p^J q^I) k \end{aligned}$$

注 1　此规则定义的四元数乘法虽然满足了乘法交换律，但由此衍生出来的交换四元数的模不满足三角不等式。例如：

$$q_1 = \frac{1+j}{2}, q_2 = \frac{1-j}{2}$$

而 $\|q_1 + q_2\| = 1 > \|q_1\| + \|q_2\| = 0$。因此，交换四元数不能直接使用以往的许多不等式，为了克服这一困难，下面的四元数神经网络将被等价转化为实数神经网络进行研究。

2.2　四元数神经网络

考虑如下的四元数神经网络：

$$\dot{\boldsymbol{q}}(t) = -\boldsymbol{C}\boldsymbol{q}(t) + \boldsymbol{A}\boldsymbol{f}(\boldsymbol{q}(t)) + \boldsymbol{B}\boldsymbol{g}(\boldsymbol{q}(t-\tau(t))) + \boldsymbol{u} \tag{3}$$

其中，$\boldsymbol{q}(t) = (q_1(t), q_2(t), \cdots, q_n(t))^{\mathrm{T}} \in \mathbb{Q}^n$ 是具有 n 个神经元的 CQVNN 在 t 时刻的状态向量。$\boldsymbol{C} = \mathrm{diag}\{c_1, c_2, \cdots, c_n\} \in \mathbb{R}^n$，$c_i > 0$ $(i = 1, 2, \cdots, n)$，是自反馈连接矩阵，$\boldsymbol{A} = \boldsymbol{A}^R + \boldsymbol{A}^I i + \boldsymbol{A}^J j + \boldsymbol{A}^K k \in \mathbb{Q}^{n\times n}$ 是连接权矩阵，$\boldsymbol{B} = \boldsymbol{B}^R + \boldsymbol{B}^I i + \boldsymbol{B}^J j + \boldsymbol{B}^K k \in \mathbb{Q}^{n\times n}$ 是时滞连接权矩阵。$\boldsymbol{f}(\boldsymbol{q}(t)) = (f_1(q_1(t)), f_2(q_2(t)), \cdots, f_n(q_n(t)))^{\mathrm{T}} \in \mathbb{Q}^n$ 是交换四元数值的激活函数，$\boldsymbol{g}(\boldsymbol{q}(t-\tau(t))) = (g_1(q_1(t-\tau(t))), g_2(q_2(t-\tau(t))), \cdots, g_n(q_n(t-\tau(t))))^{\mathrm{T}} \in \mathbb{Q}^n$ 是带有时滞的交换四元数值的激活函数。$\boldsymbol{u} = \boldsymbol{u}^R + \boldsymbol{u}^I i + \boldsymbol{u}^J j + \boldsymbol{u}^K k = (u_1, u_2, \cdots, u_n)^{\mathrm{T}} \in \mathbb{Q}^n$ 是外部输入向量，$0 \leqslant \tau(t) \leqslant \tau$ $(\tau > 0)$ 是时滞项。式（3）的初始条件如下：

$$\boldsymbol{q}(s)=\boldsymbol{\varphi}(s)\in\mathbb{Q}^n, s\in[t_0-\tau,t_0]$$
$$\boldsymbol{\varphi}(s)=\boldsymbol{\varphi}^R(s)+\boldsymbol{\varphi}^I(s)i+\boldsymbol{\varphi}^J(s)j+\boldsymbol{\varphi}^K(s)k$$

其中，t_0 为式（3）表示的系统的初始时间，τ 为时滞项 $\tau(t)$ 的上界，$\boldsymbol{\varphi}(s)$ 为系统初始条件。

假设 1 设 $q(t)=q^R(t)+q^I(t)i+q^J(t)j+q^K(t)k$，$q^R(t)$、$q^I(t)$、$q^J(t)$、$q^K(t)\in\mathbb{R}$。对于所有的 $m(m=1,2,\cdots,n)$，$f_m(q(t))$ 和 $g_m(q(t-\tau(t)))$ 能够表示成：

$$f_m(q(t))=f_m^R(q^R(t))+f_m^I(q^I(t))i+f_m^J(q^J(t))j+f_m^K(q^K(t))k$$
$$g_m(q(t-\tau(t)))=g_m^R(q^R(t-\tau(t)))+g_m^I(q^I(t-\tau(t)))i+g_m^J(q^J(t-\tau(t)))j+g_m^K(q^K(t-\tau(t)))k$$

其中，$f_m^\ell(\bullet),g_m^\ell(\bullet):\mathbb{R}^n\to\mathbb{R}^n,\ell\in\{R,I,J,K\}$，满足：

$$|f_m^\ell(x^\ell)-f_m^\ell(y^\ell)|\leqslant l_m^\ell|x^\ell-y^\ell|,$$
$$|g_m^\ell(x^\ell)-g_m^\ell(y^\ell)|\leqslant \tilde{l}_m^\ell|x^\ell-y^\ell|,\forall x^\ell,y^\ell\in\mathbb{R}^n$$

其中，l_m、$\tilde{l}_m$ 是利普希茨（Lipschitz）常数，记：

$$\boldsymbol{L}^\ell=\mathrm{diag}\{l_1^\ell,l_2^\ell,\cdots,l_n^\ell\},\tilde{\boldsymbol{L}}^\ell=\mathrm{diag}\{\tilde{l}_1^\ell,\tilde{l}_2^\ell,\cdots,\tilde{l}_n^\ell\}$$

基于式（2）和假设 1，式（3）可以被重写成如下形式：

$$\begin{aligned}\dot{\boldsymbol{q}}^R(t)=&-\boldsymbol{C}\boldsymbol{q}^R(t)+\boldsymbol{A}^R\boldsymbol{f}^R(\boldsymbol{q}^R(t))-\boldsymbol{A}^I\boldsymbol{f}^I(\boldsymbol{q}^I(t))+\\&\boldsymbol{A}^J\boldsymbol{f}^J(\boldsymbol{q}^J(t))-\boldsymbol{A}^K\boldsymbol{f}^K(\boldsymbol{q}^K(t))+\\&\boldsymbol{B}^R\boldsymbol{g}^R(\boldsymbol{q}^R(t-\tau(t)))-\boldsymbol{B}^I\boldsymbol{g}^I(\boldsymbol{q}^I(t-\tau(t)))+\\&\boldsymbol{B}^J\boldsymbol{g}^J(\boldsymbol{q}^J(t-\tau(t)))-\boldsymbol{B}^K\boldsymbol{g}^K(\boldsymbol{q}^K(t-\tau(t)))+\boldsymbol{u}^R\end{aligned}\tag{4}$$

$$\begin{aligned}\dot{\boldsymbol{q}}^I(t)=&-\boldsymbol{C}\boldsymbol{q}^I(t)+\boldsymbol{A}^R\boldsymbol{f}^I(\boldsymbol{q}^I(t))+\boldsymbol{A}^I\boldsymbol{f}^R(\boldsymbol{q}^R(t))+\\&\boldsymbol{A}^J\boldsymbol{f}^K(\boldsymbol{q}^K(t))+\boldsymbol{A}^K\boldsymbol{f}^J(\boldsymbol{q}^J(t))+\\&\boldsymbol{B}^R\boldsymbol{g}^I(\boldsymbol{q}^I(t-\tau(t)))+\boldsymbol{B}^I\boldsymbol{g}^R(\boldsymbol{q}^R(t-\tau(t)))+\\&\boldsymbol{B}^J\boldsymbol{g}^K(\boldsymbol{q}^K(t-\tau(t)))+\boldsymbol{B}^K\boldsymbol{g}^J(\boldsymbol{q}^J(t-\tau(t)))+\boldsymbol{u}^I\end{aligned}\tag{5}$$

$$\begin{aligned}\dot{\boldsymbol{q}}^J(t)=&-\boldsymbol{C}\boldsymbol{q}^J(t)+\boldsymbol{A}^R\boldsymbol{f}^J(\boldsymbol{q}^J(t))-\boldsymbol{A}^I\boldsymbol{f}^K(\boldsymbol{q}^K(t))+\\&\boldsymbol{A}^J\boldsymbol{f}^R(\boldsymbol{q}^R(t))-\boldsymbol{A}^K\boldsymbol{f}^I(\boldsymbol{q}^I(t))+\\&\boldsymbol{B}^R\boldsymbol{g}^J(\boldsymbol{q}^J(t-\tau(t)))-\boldsymbol{B}^I\boldsymbol{g}^K(\boldsymbol{q}^K(t-\tau(t)))+\\&\boldsymbol{B}^J\boldsymbol{g}^R(\boldsymbol{q}^R(t-\tau(t)))-\boldsymbol{B}^K\boldsymbol{g}^I(\boldsymbol{q}^I(t-\tau(t)))+\boldsymbol{u}^J\end{aligned}\tag{6}$$

$$\begin{aligned}\dot{\boldsymbol{q}}^K(t)=&-\boldsymbol{C}\boldsymbol{q}^K(t)+\boldsymbol{A}^R\boldsymbol{f}^K(\boldsymbol{q}^K(t))+\boldsymbol{A}^I\boldsymbol{f}^J(\boldsymbol{q}^J(t))+\\&\boldsymbol{A}^J\boldsymbol{f}^I(\boldsymbol{q}^I(t))+\boldsymbol{A}^K\boldsymbol{f}^R(\boldsymbol{q}^R(t))+\\&\boldsymbol{B}^R\boldsymbol{g}^K(\boldsymbol{q}^K(t-\tau(t)))+\boldsymbol{B}^I\boldsymbol{g}^J(\boldsymbol{q}^J(t-\tau(t)))+\\&\boldsymbol{B}^J\boldsymbol{g}^I(\boldsymbol{q}^I(t-\tau(t)))+\boldsymbol{B}^K\boldsymbol{g}^R(\boldsymbol{q}^R(t-\tau(t)))+\boldsymbol{u}^K\end{aligned}\tag{7}$$

根据式（4）～式（7），可以得到：

$$\dot{\boldsymbol{Q}}(t)=-\hat{\boldsymbol{C}}\boldsymbol{Q}(t)+\hat{\boldsymbol{A}}\hat{\boldsymbol{F}}(\boldsymbol{Q}(t))+\hat{\boldsymbol{B}}\hat{\boldsymbol{G}}(\boldsymbol{Q}(t-\tau(t)))+\boldsymbol{U}\tag{8}$$

其中，

$$\hat{\boldsymbol{A}}=\begin{pmatrix}\boldsymbol{A}^R&-\boldsymbol{A}^I&\boldsymbol{A}^J&-\boldsymbol{A}^K\\\boldsymbol{A}^I&\boldsymbol{A}^R&\boldsymbol{A}^K&\boldsymbol{A}^J\\\boldsymbol{A}^J&-\boldsymbol{A}^K&\boldsymbol{A}^R&-\boldsymbol{A}^I\\\boldsymbol{A}^K&\boldsymbol{A}^J&\boldsymbol{A}^I&\boldsymbol{A}^R\end{pmatrix}$$

$$\hat{\boldsymbol{B}}=\begin{pmatrix}\boldsymbol{B}^R&-\boldsymbol{B}^I&\boldsymbol{B}^J&-\boldsymbol{B}^K\\\boldsymbol{B}^I&\boldsymbol{B}^R&\boldsymbol{B}^K&\boldsymbol{B}^J\\\boldsymbol{B}^J&-\boldsymbol{B}^K&\boldsymbol{B}^R&-\boldsymbol{B}^I\\\boldsymbol{B}^K&\boldsymbol{B}^J&\boldsymbol{B}^I&\boldsymbol{B}^R\end{pmatrix}$$

$$\boldsymbol{Q}(t)=((\boldsymbol{q}^R(t))^{\mathrm{T}},(\boldsymbol{q}^I(t))^{\mathrm{T}},(\boldsymbol{q}^J(t))^{\mathrm{T}},(\boldsymbol{q}^K(t))^{\mathrm{T}})^{\mathrm{T}}$$
$$\boldsymbol{U}=((\boldsymbol{u}^R)^{\mathrm{T}},(\boldsymbol{u}^I)^{\mathrm{T}},(\boldsymbol{u}^J)^{\mathrm{T}},(\boldsymbol{u}^K)^{\mathrm{T}})^{\mathrm{T}}$$
$$\hat{\boldsymbol{C}}=\mathrm{diag}\{\boldsymbol{C},\boldsymbol{C},\boldsymbol{C},\boldsymbol{C}\}$$
$$\hat{\boldsymbol{F}}=((\boldsymbol{f}^R)^{\mathrm{T}},(\boldsymbol{f}^I)^{\mathrm{T}},(\boldsymbol{f}^J)^{\mathrm{T}},(\boldsymbol{f}^J)^{\mathrm{T}})^{\mathrm{T}}$$
$$\hat{\boldsymbol{G}}=((\boldsymbol{g}^R)^{\mathrm{T}},(\boldsymbol{g}^I)^{\mathrm{T}},(\boldsymbol{g}^J)^{\mathrm{T}},(\boldsymbol{g}^J)^{\mathrm{T}})^{\mathrm{T}}$$

通过假设 1，可以得到：

$$|\hat{\boldsymbol{F}}(\boldsymbol{Q}^\ell)-\hat{\boldsymbol{F}}(\bar{\boldsymbol{Q}}^\ell)|\leqslant\boldsymbol{L}|\boldsymbol{Q}^\ell-\bar{\boldsymbol{Q}}^\ell|$$
$$|\hat{\boldsymbol{G}}(\boldsymbol{Q}^\ell)-\hat{\boldsymbol{G}}(\bar{\boldsymbol{Q}}^\ell)|\leqslant\tilde{\boldsymbol{L}}|\boldsymbol{Q}^\ell-\bar{\boldsymbol{Q}}^\ell|,\forall\boldsymbol{Q}^\ell,\bar{\boldsymbol{Q}}^\ell\in\mathbb{R}^{4n}$$

其中，$\boldsymbol{L}=\mathrm{diag}\{\boldsymbol{L}^R,\boldsymbol{L}^I,\boldsymbol{L}^J,\boldsymbol{L}^K\}$，$\tilde{\boldsymbol{L}}=\mathrm{diag}\{\tilde{\boldsymbol{L}}^R,\tilde{\boldsymbol{L}}^I,\tilde{\boldsymbol{L}}^J,\tilde{\boldsymbol{L}}^K\}$。

定义 1 若常数向量 $\tilde{\boldsymbol{q}}=(\tilde{q}_1,\tilde{q}_2,\cdots,\tilde{q}_n)^{\mathrm{T}}$ 满足 $-\boldsymbol{C}\tilde{\boldsymbol{q}}+\boldsymbol{A}\boldsymbol{f}(\tilde{\boldsymbol{q}})+\boldsymbol{B}\boldsymbol{g}(\tilde{\boldsymbol{q}})+\boldsymbol{u}=0$，则称常数向量 $\tilde{\boldsymbol{q}}$ 是交换四元数神经网络（式（3））的平衡点。

为了保证式（7）表示的系统有唯一平衡点，需要做如下假设。

假设 2 $\hat{\boldsymbol{C}}-|\hat{\boldsymbol{A}}|\boldsymbol{L}-|\hat{\boldsymbol{B}}|\tilde{\boldsymbol{L}}$ 是一个非奇异 M 矩阵。

假设 3 设 $\boldsymbol{K}=\hat{\boldsymbol{C}}-|\hat{\boldsymbol{A}}|\boldsymbol{L}-|\hat{\boldsymbol{B}}|\tilde{\boldsymbol{L}}$，$\boldsymbol{K}+\boldsymbol{K}^{\mathrm{T}}$ 是一个非奇异矩阵。

定理 1 如果假设 1、假设 2 成立，则式（8）表示的系统存在唯一的平衡点。

定理 2 如果假设 1、假设 3 成立，则式（8）表示的系统存在唯一的平衡点。

证明 通过文献[11]中的定理 1 与定理 2，可以直接得到本文的定理 1 和定理 2。

为了更加方便地研究式（8）表示的实数系统，将其系统平衡点平移至原点。假设 $\tilde{\boldsymbol{Q}}=((\tilde{\boldsymbol{q}}^R)^{\mathrm{T}},(\tilde{\boldsymbol{q}}^I)^{\mathrm{T}},(\tilde{\boldsymbol{q}}^J)^{\mathrm{T}},(\tilde{\boldsymbol{q}}^K)^{\mathrm{T}})^{\mathrm{T}}$ 是式（8）表示的系统的平衡点，做变换 $\bar{\boldsymbol{Q}}(t)=\boldsymbol{Q}(t)-\tilde{\boldsymbol{Q}}$，则式（8）变为下面的向量形式：

$$\dot{\bar{\boldsymbol{Q}}}(t)=-\hat{\boldsymbol{C}}\bar{\boldsymbol{Q}}(t)+\hat{\boldsymbol{A}}\boldsymbol{F}(\bar{\boldsymbol{Q}}(t))+\hat{\boldsymbol{B}}\boldsymbol{G}(\bar{\boldsymbol{Q}}(t-\tau(t)))\quad(9)$$

其中，$\boldsymbol{F}(\bar{\boldsymbol{Q}}(t))=\hat{\boldsymbol{F}}(\bar{\boldsymbol{Q}}(t)+\tilde{\boldsymbol{Q}})-\hat{\boldsymbol{F}}(\tilde{\boldsymbol{Q}})$, $\boldsymbol{G}(\bar{\boldsymbol{Q}}(t-\tau(t)))=\hat{\boldsymbol{G}}(\bar{\boldsymbol{Q}}(t-\tau(t))+\tilde{\boldsymbol{Q}})-\hat{\boldsymbol{G}}(\tilde{\boldsymbol{Q}})$。

显然，式（9）表示的实值系统原点的稳定性与式（8）表示的实值系统及交换四元数神经网络（如式（3）所示）的平衡点的稳定性是等价的。

下面对式（9）的渐近稳定性做一个系统的定义。

定义 2　若对于任意给定的 $\varepsilon>0$，都能找到正数 $\delta=\delta(\varepsilon,t_0)$，使得当 $\|x_0\|<\delta$ 和 $t>t_0$ 时，式（9）的解 $\bar{\boldsymbol{Q}}(t)=\bar{\boldsymbol{Q}}(t,t_0,x_0)$ 满足 $\|\bar{\boldsymbol{Q}}(t,t_0,x_0)\|<\varepsilon$，则称式（9）的零解是稳定的。

定义 3　设 U 是 $\mathbb{R}^n$ 中包含坐标原点的一个开区域，若对于所有的 $x_0\in U$ 和任意给定的 $\varepsilon>0$，都能找到正数 $T=T(\varepsilon,t_0,x_0)$，使得当 $t>t_0+T$ 时，式（9）的解 $\bar{\boldsymbol{Q}}(t)=\bar{\boldsymbol{Q}}(t,t_0,x_0)$ 满足 $\|\bar{\boldsymbol{Q}}(t,t_0,x_0)\|<\varepsilon$，则称式（9）的零解是吸引的，称 U 是式（9）的零解的一个吸引域。

定义 4　若式（9）的零解既是稳定的又是吸引的，则称式（9）的零解是渐近稳定的；若式（9）的零解的吸引域是整个 $\mathbb{R}^n$，则称式（9）的零解是全局渐近稳定的。

注 2　经过变换后，式（9）的平衡点平移至原点，此时称系统的解为零解。利用极限的形式可以给出式（9）的零解稳定和吸引的直观描述为：

当对于一切 $t>t_0$，$\lim\limits_{x_0\to 0}x(t,t_0,x_0)=0$ 时，称式（9）的零解是稳定的；

当对于一切的 $x_0\in U$，$\lim\limits_{t\to\infty}x(t,t_0,x_0)=0$ 时，称式（9）的零解是吸引的。

引理 1　对于任意 $\boldsymbol{a},\boldsymbol{b}\in\mathbb{R}^n$，如果 $\boldsymbol{P}$ 是一个正定的对称矩阵，则：

$$\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}+\boldsymbol{b}^{\mathrm{T}}\boldsymbol{a}\leqslant\boldsymbol{a}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{a}+\boldsymbol{b}^{\mathrm{T}}\boldsymbol{P}^{-1}\boldsymbol{b}$$

证明　根据文献[12]中的引理 2，可以直接得到这个结果。

引理 2[12]　当且仅当下面的两个式子之一成立：

(i)　$\boldsymbol{S}_{22}<0$ 和 $\boldsymbol{S}_{11}-\boldsymbol{S}_{12}\boldsymbol{S}_{22}^{-1}\boldsymbol{S}_{21}<0$,

(ii)　$\boldsymbol{S}_{11}<0$ 和 $\boldsymbol{S}_{22}-\boldsymbol{S}_{21}\boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{12}<0$。

则有实对称矩阵：

$$\boldsymbol{S}=\begin{pmatrix}\boldsymbol{S}_{11} & \boldsymbol{S}_{12}\\ \boldsymbol{S}_{21} & \boldsymbol{S}_{22}\end{pmatrix}<0$$

其中，$\boldsymbol{S}_{11}^{\mathrm{T}}=\boldsymbol{S}_{11}$、$\boldsymbol{S}_{12}^{\mathrm{T}}=\boldsymbol{S}_{21}$ 和 $\boldsymbol{S}_{22}^{\mathrm{T}}=\boldsymbol{S}_{22}$。

引理 3[13]　若在原点的领域 U 内，存在正定（负定）函数 $V(t)$ 使得 $\dot{V}(t)$ 是负定（正定）的，那么式（9）的零解是渐近稳定的。

注 3　无论是在深度学习领域还是在神经网络动力学行为研究领域，考虑比较多的都是实数神经网络或者复数值神经网络，这些神经网络在处理三维或者四维等高维数据时具有一定的局限性。于是，近几年出现了基于 Hamilton 规则的四元数神经网络，这种神经网络对高维数据的处理具有一定的优势。但是，基于 Hamilton 规则的四元数不满足乘法交换律，本文给出了可交换的四元数神经网络模型，一定程度上解决了基于 Hamilton 规则的四元数神经网络具有的局限性。

3　渐进稳定性

本节将基于李雅普诺夫（Lyapunov）稳定性理论给出式（8）表示的系统的唯一平衡点渐近稳定的充分性判据。

定理 3　在假设 1、2 的条件下，如果存在 4 个正定对角矩阵 $\boldsymbol{P}_1$、$\boldsymbol{P}_2$、$\boldsymbol{\Gamma}_1$、$\boldsymbol{\Gamma}_2$ 使得如下线性矩阵不等式（linear matrix inequality，LMI）条件成立：

$$\begin{pmatrix}-\hat{\boldsymbol{C}}\boldsymbol{P}_1-\boldsymbol{P}_1\hat{\boldsymbol{C}}+\boldsymbol{P}_2+\boldsymbol{L}\boldsymbol{\Gamma}_1\boldsymbol{L} & \boldsymbol{P}_1\hat{\boldsymbol{A}} & \boldsymbol{P}_1\hat{\boldsymbol{B}}\\ \hat{\boldsymbol{A}}^T\boldsymbol{P}_1 & -\boldsymbol{\Gamma}_1 & \boldsymbol{0}\\ \hat{\boldsymbol{B}}^T\boldsymbol{P}_1 & \boldsymbol{0} & -\boldsymbol{\Gamma}_2\end{pmatrix}<0\quad(10)$$

$$\tilde{\boldsymbol{L}}\boldsymbol{\Gamma}_2\tilde{\boldsymbol{L}}-(1-\tau)\boldsymbol{P}_2<0\quad(11)$$

则式（9）表示系统的唯一平衡点是渐近稳定的。

证明：构造如下 Lyapunov 函数：

$$V(t)=\bar{\boldsymbol{Q}}^{\mathrm{T}}(t)\boldsymbol{P}_1\bar{\boldsymbol{Q}}(t)+\int_{t-\tau(t)}^{t}\bar{\boldsymbol{Q}}^{\mathrm{T}}(s)\boldsymbol{P}_2\bar{\boldsymbol{Q}}(s)\mathrm{d}s$$

根据假设 1 及引理 1 计算 $V(t)$ 的导数：

$$\begin{aligned}\dot{V}(t)=&\dot{\bar{\boldsymbol{Q}}}^{\mathrm{T}}(t)\boldsymbol{P}_1\bar{\boldsymbol{Q}}(t)+\bar{\boldsymbol{Q}}^{\mathrm{T}}(t)\boldsymbol{P}_1\dot{\bar{\boldsymbol{Q}}}(t)+\\&\bar{\boldsymbol{Q}}^{\mathrm{T}}(t)\boldsymbol{P}_2\bar{\boldsymbol{Q}}(t)-(1-\dot{\tau}(t))\times\\&\bar{\boldsymbol{Q}}^{\mathrm{T}}(t-\tau(t))\boldsymbol{P}_2\bar{\boldsymbol{Q}}(t-\tau(t))=\\&\bar{\boldsymbol{Q}}^{\mathrm{T}}(t)(-\hat{\boldsymbol{C}}\boldsymbol{P}_1-\boldsymbol{P}_1\hat{\boldsymbol{C}}+\boldsymbol{P}_2)\bar{\boldsymbol{Q}}(t)+\\&\boldsymbol{F}^{\mathrm{T}}(\bar{\boldsymbol{Q}}(t))\hat{\boldsymbol{A}}^T\boldsymbol{P}_1\bar{\boldsymbol{Q}}(t)+\bar{\boldsymbol{Q}}^{\mathrm{T}}(t)\boldsymbol{P}_1\hat{\boldsymbol{A}}\boldsymbol{F}(\bar{\boldsymbol{Q}}(t))+\\&\boldsymbol{G}^{\mathrm{T}}(\bar{\boldsymbol{Q}}(t-\tau(t)))\hat{\boldsymbol{B}}^{\mathrm{T}}\boldsymbol{P}_1\bar{\boldsymbol{Q}}(t)+\\&\bar{\boldsymbol{Q}}(t)\boldsymbol{P}_1\hat{\boldsymbol{B}}\boldsymbol{G}^{\mathrm{T}}(\bar{\boldsymbol{Q}}(t-\tau(t)))-\\&(1-\dot{\tau}(t))\bar{\boldsymbol{Q}}^{\mathrm{T}}(t-\tau(t))\boldsymbol{P}_2\bar{\boldsymbol{Q}}(t-\tau(t))\leqslant\\&\bar{\boldsymbol{Q}}^{\mathrm{T}}(t)(-\hat{\boldsymbol{C}}\boldsymbol{P}_1-\boldsymbol{P}_1\hat{\boldsymbol{C}}+\boldsymbol{P}_2)\bar{\boldsymbol{Q}}(t)+\end{aligned}$$

$$\begin{aligned}&\boldsymbol{F}^{\mathrm{T}}(\bar{\boldsymbol{Q}}(t))\boldsymbol{\Gamma}_1\boldsymbol{F}(\bar{\boldsymbol{Q}}(t))+\\&\bar{\boldsymbol{Q}}^{\mathrm{T}}(t)\boldsymbol{P}_1\hat{\boldsymbol{A}}\boldsymbol{\Gamma}_1^{-1}\hat{\boldsymbol{A}}^{\mathrm{T}}\boldsymbol{P}_1\bar{\boldsymbol{Q}}(t)+\\&\boldsymbol{G}^{\mathrm{T}}(\bar{\boldsymbol{Q}}(t-\tau(t)))\boldsymbol{\Gamma}_2\boldsymbol{G}(\bar{\boldsymbol{Q}}(t-\tau(t)))+\\&\bar{\boldsymbol{Q}}^{\mathrm{T}}(t)\boldsymbol{P}_1\hat{\boldsymbol{B}}\boldsymbol{\Gamma}_1^{-1}\hat{\boldsymbol{B}}^{\mathrm{T}}\boldsymbol{P}_1\bar{\boldsymbol{Q}}(t)-(1-\tau)\times\\&\bar{\boldsymbol{Q}}^{\mathrm{T}}(t-\tau(t))\boldsymbol{P}_2\bar{\boldsymbol{Q}}(t-\tau(t))\leqslant\\&\bar{\boldsymbol{Q}}^{\mathrm{T}}(t)(-\hat{\boldsymbol{C}}\boldsymbol{P}_1-\boldsymbol{P}_1\hat{\boldsymbol{C}}+\boldsymbol{P}_2+\\&\boldsymbol{P}_1\hat{\boldsymbol{A}}\boldsymbol{\Gamma}_1^{-1}\hat{\boldsymbol{A}}^{\mathrm{T}}\boldsymbol{P}+\boldsymbol{P}_1\hat{\boldsymbol{B}}\boldsymbol{\Gamma}_1^{-1}\hat{\boldsymbol{B}}^{\mathrm{T}}\boldsymbol{P}_1)\bar{\boldsymbol{Q}}(t)+\\&\bar{\boldsymbol{Q}}^{\mathrm{T}}(t-\tau(t))(\tilde{\boldsymbol{L}}\boldsymbol{\Gamma}_2\tilde{\boldsymbol{L}}-(1-\tau)\boldsymbol{P}_2)\bar{\boldsymbol{Q}}(t-\tau(t))\end{aligned}$$

根据引理 2 和式（10）得出 $-\hat{\boldsymbol{C}}\boldsymbol{P}_1-\boldsymbol{P}_1\hat{\boldsymbol{C}}+\boldsymbol{P}_2+\boldsymbol{P}_1\hat{\boldsymbol{A}}\boldsymbol{\Gamma}_1^{-1}\hat{\boldsymbol{A}}^{\mathrm{T}}\boldsymbol{P}+\boldsymbol{P}_1\hat{\boldsymbol{B}}\boldsymbol{\Gamma}_1^{-1}\hat{\boldsymbol{B}}^{\mathrm{T}}\boldsymbol{P}<0$，结合式（11），得到 $V(t)<0$。又因为 $V(t)$ 是径向无界的，根据一般的 Lyapunov 稳定性理论（引理 3）得知式（9）表示的系统的唯一平衡点是渐近稳定的。

4　案例分析

为了验证定理 3 得到的稳定性判据的有效性，本节将选择一个符合定理 3 的系统案例，并通过 Matlab 绘制其解曲线并观察解的轨迹，进而判定本文所得结果的有效性。另外，考虑到现实中噪声的影响，在案例 2 中对系统的连接权以及系统的外部输入项添加一个正态分布的（0,1）噪声，然后观察分析系统解曲线的性态。

（1）案例 1

考虑如下二维交换四元数神经网络：

$$\dot{\boldsymbol{q}}(t)=-\boldsymbol{C}\boldsymbol{q}(t)+\boldsymbol{A}\boldsymbol{f}(\boldsymbol{q}(t))+\boldsymbol{B}\boldsymbol{g}(\boldsymbol{q}(t-\tau(t)))+\boldsymbol{u}\tag{12}$$

其中，$\boldsymbol{A}=(a_{ij})_{2\times2}$, $\boldsymbol{B}=(b_{ij})_{2\times2}$,

$$\begin{aligned}&a_{11}=-0.1-0.4i-0.1j-0.5k,\\&a_{12}=-0.4+0.2i-0.4j+0.2k,\\&a_{21}=-0.3+0.1i-0.2j+0.1k,\\&a_{22}=-0.6-0.1i-0.2j-0.1k,\\&b_{11}=0.2-0.1i+0.1j+0.1k,\\&b_{12}=0.2+0.1i-0.2j-0.3k,\\&b_{21}=-0.1+0.1i-0.1j-0.2k,\\&b_{22}=0.3-0.3i+0.1j+0.2k,\\&\boldsymbol{C}=\mathrm{diag}\{4,4\},\tau(t)=0.5\sin^2(t),\\&\boldsymbol{f}(\boldsymbol{q}(t))=\tanh(\boldsymbol{q}^R(t))+\tanh(\boldsymbol{q}^I(t))i+\\&\qquad\tanh(\boldsymbol{q}^J(t))j+\tanh(\boldsymbol{q}^K(t))k,\\&\boldsymbol{g}(\boldsymbol{q}(t))=\tanh(\boldsymbol{q}^R(t))+\tanh(\boldsymbol{q}^I(t))i+\\&\qquad\tanh(\boldsymbol{q}^J(t))j+\tanh(\boldsymbol{q}^K(t))k。\end{aligned}$$

取 $\boldsymbol{L}^{\ell}=\tilde{\boldsymbol{L}}^{\ell}=\mathrm{diag}\{1,1\},\ell\in\{R,I,J,K\}$，容易验证 $\hat{\boldsymbol{C}}-|\hat{\boldsymbol{A}}|\boldsymbol{L}-|\hat{\boldsymbol{B}}|\tilde{\boldsymbol{L}}$ 是 L 矩阵，经计算得到 $\hat{\boldsymbol{C}}-|\hat{\boldsymbol{A}}|\boldsymbol{L}-|\hat{\boldsymbol{B}}|\tilde{\boldsymbol{L}}$ 的特征值为 0.694、4.669、4.000、3.900、3.806、3.400、3.432、3.300，它们全具有正实部，则 $\hat{\boldsymbol{C}}-|\hat{\boldsymbol{A}}|\boldsymbol{L}-|\hat{\boldsymbol{B}}|\tilde{\boldsymbol{L}}$ 是非奇异 M 矩阵。通过 Matlab 中的 YALMIP 工具包，得到正定的对角矩阵 $\boldsymbol{P}_1$、$\boldsymbol{P}_2$、$\boldsymbol{\Gamma}_1$、$\boldsymbol{\Gamma}_2$，它们满足式（10）和式（11）。

$$\begin{aligned}&\boldsymbol{P}_1=\mathrm{diag}\{\boldsymbol{P}_{11},\boldsymbol{P}_{11},\boldsymbol{P}_{11},\boldsymbol{P}_{11}\}\\&\boldsymbol{P}_2=\mathrm{diag}\{\boldsymbol{P}_{22},\boldsymbol{P}_{22},\boldsymbol{P}_{22},\boldsymbol{P}_{22}\}\\&\boldsymbol{\Gamma}_1=\mathrm{diag}\{\boldsymbol{\Gamma}_{11},\boldsymbol{\Gamma}_{11},\boldsymbol{\Gamma}_{11},\boldsymbol{\Gamma}_{11}\}\\&\boldsymbol{\Gamma}_2=\mathrm{diag}\{\boldsymbol{\Gamma}_{22},\boldsymbol{\Gamma}_{22},\boldsymbol{\Gamma}_{22},\boldsymbol{\Gamma}_{22}\}\\&\boldsymbol{P}_{11}=\begin{pmatrix}3.5668&0\\0&3.6169\end{pmatrix}\\&\boldsymbol{P}_{22}=\begin{pmatrix}11.9318&0\\0&12.0776\end{pmatrix}\\&\boldsymbol{\Gamma}_{11}=\begin{pmatrix}8.3862&0\\0&8.4712\end{pmatrix}\\&\boldsymbol{\Gamma}_{22}=\begin{pmatrix}4.8406&0\\0&4.8649\end{pmatrix}\end{aligned}$$

通过定理 3，可以得到式（12）表示的系统是渐近稳定的。这里选取系统的初始值为：

$$\boldsymbol{q}(s)=\begin{pmatrix}1.2-1.3i+1.4j+1.9k\\-1.6+1.8i-15j-1.6k\end{pmatrix}$$

其中，$s\in[-0.5,0]$，图 1～图 4 展示了所给系统的 4 个部分的状态轨迹，这里的状态轨迹是指系统的解曲线，从这 4 幅图中可以看到每个神经元都收敛到稳定状态，收敛点为 $\tilde{\boldsymbol{q}}=(\tilde{q}_{ij})_{2\times1}$，其中 $\tilde{q}_{11}=0.439+0.186i+0.383j+0.287k$，$\tilde{q}_{21}=0.324+0.362i+0.274j+0.386k$。

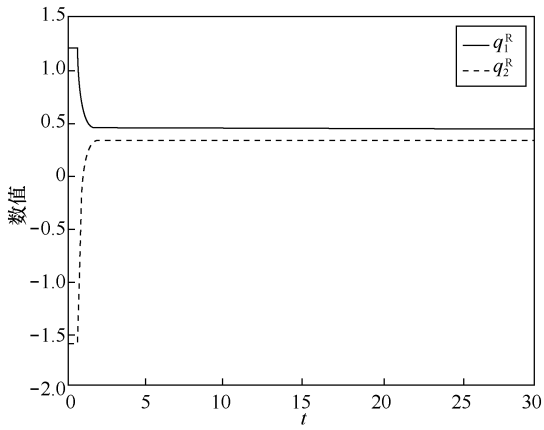

图 1　二维四元数神经网络第一部分的状态轨迹

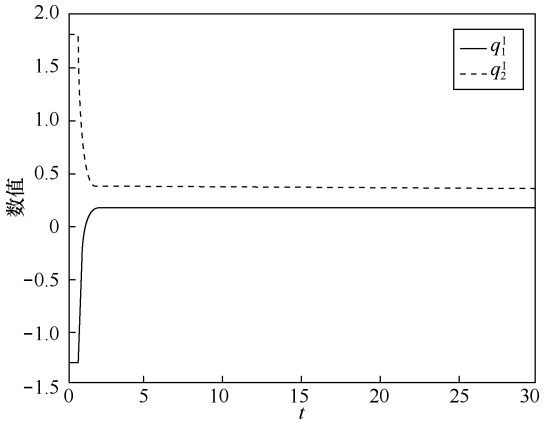

图 2　二维四元数神经网络第二部分的状态轨迹

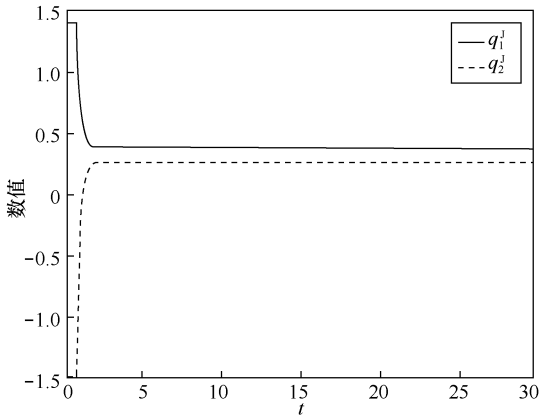

图 3　二维四元数神经网络第三部分的状态轨迹

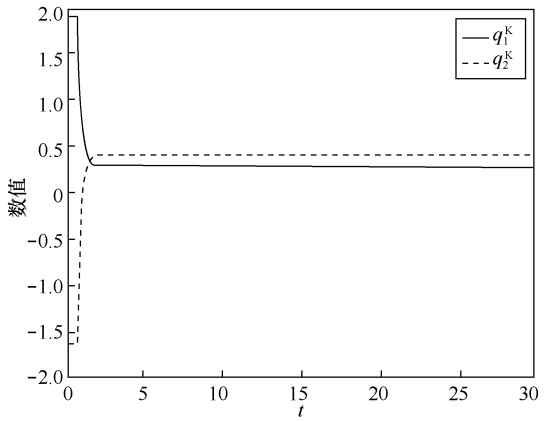

图 4　二维四元数神经网络第四部分的状态轨迹

图 1～图 8 中 q_1 和 q_2 分别为式（12）所示系统的第一维和第二维的解，t 表示 q_1 和 q_2 的自变量。

（2）案例 2

对案例 1 中涉及的系统的连接权参数 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 以及外部输入 $\boldsymbol{u}$ 添加一个正态分布的（0,1）噪声，通过 Matlab 的数值仿真结果如图 5～图 8 所示，结果显示每一个神经元都收敛于平衡点。可见系统在受到一定噪声干扰时，其稳定性不会受到影响。

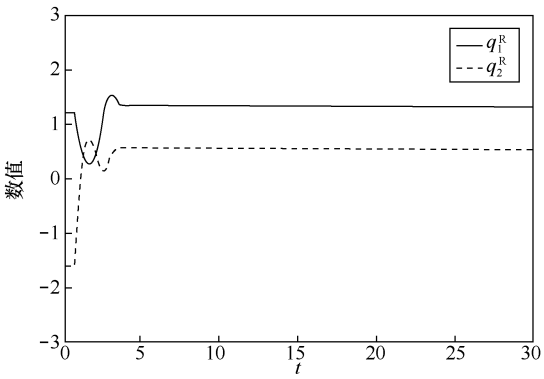

图 5　二维四元数神经网络第一部分受噪声干扰后的状态轨迹

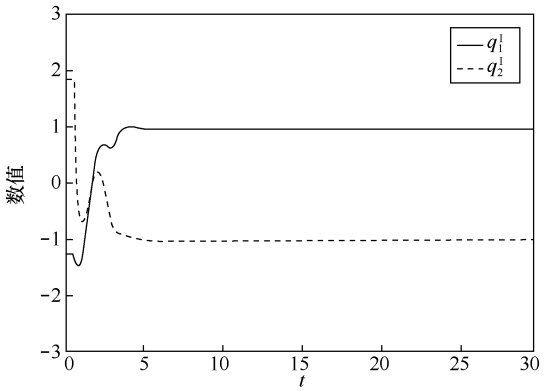

图 6　二维四元数神经网络第二部分受噪声干扰后的状态轨迹

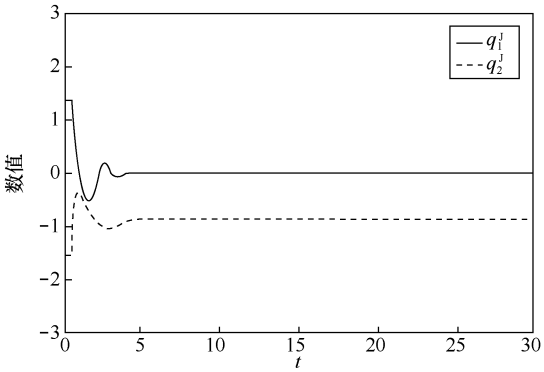

图 7　二维四元数神经网络第三部分受噪声干扰后的状态轨迹

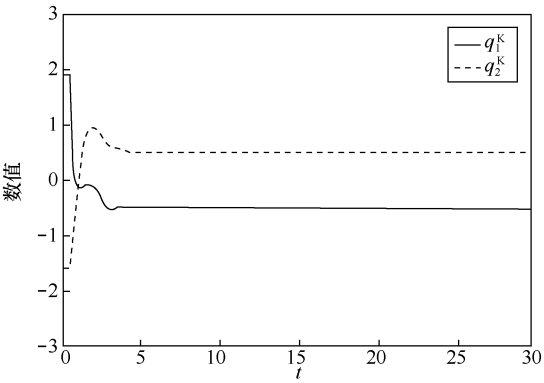

图 8　二维四元数神经网络第四部分受噪声干扰后的状态轨迹

注 4 按照数据类型，可将神经网络分为实值神经网络、复值神经网络和四元数神经网络。实值神经网络的应用十分广泛，但是它无法高效地处理诸如二维仿射变换或者三维仿射变换等几何变换问题。为了解决这一问题，学者们发现复值神经网络能够将二维空间中的一个点数据建模成一个单独的整体，而不是两个数据单元的集合，这就使得复值神经网络在处理二维数据的仿射变换时，比实值神经网络更为高效快捷。然而，在实际生活中，常常需要处理三维数据，例如，与图像有关的应用、方向预测以及机器人操控等。近年来，学者们发现采用四元数可以直接对高维数据进行编码处理，大幅提升了处理效率，尤其在处理三维空间变换时非常高效简洁。目前四元数神经网络已经在卫星的姿态控制、彩色图像压缩、计算机图像学、彩色微光夜视以及机器人运动控制等诸多领域展现了良好的应用前景[14-17]。因此，关于四元数神经网络稳定性的研究是十分有意义的。然而，目前研究的四元数神经网络不满足乘法交换律，即存在一定的局限性。本文提出的可交换四元数神经网络在一定程度上解决了这一局限性。目前，关于本文所涉及的交换四元数神经网络的渐近稳定性研究较少，因此本文研究具有一定的创新性。

5 结束语

本文对交换四元数神经网络的渐近稳定性进行了研究。首先介绍了交换四元数的一些基本概念及运算，然后提出了交换四元数的神经网络。考虑到交换四元数具有 3 个共轭这一事实所带来的各种问题，文中将交换四元数神经网络分解成 4 个实值神经网络，通过不等式理论、Lyapunov 稳定性理论以及拓扑度理论，得到了一些交换四元数神经网络平衡点存在且唯一的充分条件以及平衡点渐近稳定的 LMI 条件。最后，提供了两个案例来证明所提出的 LMI 条件的有效性。

参考文献：

[1] HOPFIELD J J. Neural networks and physical systems with emergent collective computational abilities[J]. Proceedings of the National Academy of Sciences, 1982, 79(8): 2554-2558.

[2] ZENG Z, WANG J. Design and analysis of high-capacity associative memories based on a class of discrete-time recurrent neural networks[J]. IEEE Transactions on Systems Man & Cybernetics Part B Cybernetics: A Publication of the IEEE Systems Man & Cybernetics Society, 2008, 38(6):1525-1536.

[3] TANAKA G, AIHARA K. Complex-valued multistate associative memory with nonlinear multilevel functions for gray-level image reconstruction[J]. IEEE Transactions on Neural Networks, 2009, 20(9): 1463-1473.

[4] CHEN Y H, FANG S C. Neurocomputing with time delay analysis for solving convex quadratic programming problems[J]. IEEE Transactions on Neural Networks, 2000, 11(1): 230-240.

[5] AMIN M F, MURASE K. Single-layered complex-valued neural network for real-valued classification problems[J]. Neurocomputing, 2009, 72(4-6): 945-955.

[6] SONG Q, YAN H, ZHAO Z, et al. Global exponential stability of complex-valued neural networks with both time-varying delays and impulsive effects[J]. Neural Networks, 2016, 79:108-116.

[7] ZHANG H, YANG F, LIU X, et al. Stability analysis for neural networks with time-varying delay based on quadratic convex combination[J]. IEEE Transactions on Neural Networks and Learning Systems, 2013, 24(4): 513-521.

[8] SONG Q, CHEN X. Multistability analysis of quaternion-valued neural networks with time delays[J]. IEEE Transactions on Neural Networks and Learning Systems, 2018, 29(11): 5430-5440.

[9] ZENG R, WU J, SHAO Z, et al. Color image classification via quaternion principal component analysis network[J]. Neurocomputing, 2016, 216: 416-428.

[10] CHEN B, SHU H, COATRIEUX G, et al. Color image analysis by quaternion-type moments[J]. Journal of Mathematical Imaging and Vision, 2015, 51(1): 124-144.

[11] CAO J D, WANG J. Global asymptotic stability of a general class of recurrent neural networks with time-varying delays[J]. IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications, 2003, 50(1): 34-44.

[12] CHEN X F, LI Z S, SONG Q K, et al. Robust stability analysis of quaternion-valued neural networks with time delays and parameter uncertainties[J]. Neural Networks, 2017, 91: 55-65.

[13] 马知恩, 周义仓, 李承治. 常微分方程定性与稳定性方法[M]. 北京: 科学出版社, 2015.
MA Z E, ZHOU Y C, LI C Z. Qualitative and stability methods for ordinary differential equations[M]. Beijing: Science Press, 2015.

[14] ISOKAWA T, KUSAKABE T, MATSUI N, et al. Quaternion neural network and its application[C]// International Conference on Knowledge-Based and Intelligent Information and Engineering Systems. Heidelberg: Springer, 2003: 318-324.

[15] GUPTA S. Linear quaternion equations with application to spacecraft attitude propagation[C]// 1998 IEEE Aerospace Conference. Piscataway: IEEE Press, 1998, 1: 69-76.

[16] KUSAMICHI H, ISOKAWA T, MATSUI N, et al. A new scheme for

color night vision by quaternion neural network[C]// The 2nd International Conference on Autonomous Robots and Agents. Piscataway: IEEE Press, 2004: 1315.

[17] MATSUI N, ISOKAWA T, KUSAMICHI H, et al. Quaternion neural network with geometrical operators[J]. Journal of Intelligent & Fuzzy Systems, 2004, 15(3, 4): 149-164.

[作者简介]

林东源（1996− ），男，重庆交通大学数学与统计学院硕士生，主要研究方向为四元数神经网络的动力学行为。

陈晓丰（1980− ），男，博士，重庆交通大学数学与统计学院副教授、硕士生导师，主要研究方向为复数及四元数神经网络稳定性理论。

孙文涛（1979− ），男，西藏大学理学院讲师，主要研究方向为四元数神经网络稳定性理论及应用。

李伟凯（1994− ），男，南京航空航天大学计算机科学与技术学院博士生，主要研究方向为机器学习、迁移学习、领域适配。

夏砚楠（1997− ），男，重庆交通大学数学与统计学院硕士生，主要研究方向为分数阶四元数神经网络的动力学行为。