

面向时变复数西尔维斯特方程的有限时间神经网络研究^①高 畅^② 孔 颖^③

(浙江科技学院信息与电子工程学院 杭州 310023)

摘 要 针对时变复数西尔维斯特方程的实时求解问题,提出了两种有限时间神经网络 (FTNN) 模型。该方案基于张神经网络 (ZNN) 在实数域中的动力学方法,设计面向复数域的神经动力学方程。针对动力学方程中非线性激励函数的数值计算问题,应用两种等价的处理方法。第一种方法是处理复数输入的实部与虚部,第二种方法是处理复数输入的模数。通过使用有限值激励函数加快 FTNN 模型的求解速度,进一步提高了模型的收敛速度和计算精度。实验结果表明,在求解时变复数西尔维斯特方程时,相比于传统的周期神经网络求解法,所提出的网络模型具有更好的收敛性和稳定性。

关键词 时变复数西尔维斯特方程;有限时间神经网络 (FTNN);有限值激励函数;收敛速度;计算精度

0 引言

矩阵方程求解广泛应用于科学研究和实际工程领域中。其中,西尔维斯特方程的求解是工程领域和数字应用中的一个主要问题,例如图像处理^[1]、机器人应用^[2]和特征值分配^[3]。线性矩阵方程 $A(t)X(t) - X(t)B(t) + C(t) = 0$ 被称为西尔维斯特方程^[4-5],它与动力学系统理论紧密相关。该方程的系数一般情况下为时变矩阵 $A(t)$ 、 $B(t)$ 和 $C(t)$, $X(t)$ 为所求解的未知矩阵。目前,求解西尔维斯特方程主要有两种方法。第一种方法源于串行处理,例如传统的数值算法^[6]。该方法通常用于求解静态情况下的西尔维斯特方程,而不适用于时变情况。另外,由于串行计算会导致采样率过高并且使得每个计算周期更为复杂,因此无法实时求解规模较大的时变问题以及更为复杂的时变复数问题。第二种方法源于并行处理,例如一些神经网络算法^[7-8],它们可以被用来求解时变情况下的西尔

维斯特方程(简称时变西尔维斯特方程)。由于神经网络算法拥有并行计算的能力而且容易在电子硬件中实现,因此具有较高的性能,可以实时计算大量数据。基于传统梯度法的梯度神经网络 (gradient neural network, GNN) 是一种典型的神经动力学方法,可以被用来求解西尔维斯特方程。但是,该方法对于实时问题的求解基本上都是考虑定常(即静态)情况的,或将时变问题借助于短时不变性假设而近似为定常问题去处理。换句话说,在求解过程中,由时变因素引起的影响经常被忽略。因此在求解时变问题时,它的估计误差可能不会收敛到 0,故该方法在求解时变西尔维斯特方程时,计算精度不高。

针对时变问题的求解,文献[9,10]提出了一类新型的递归神经网络 (zeroing neural network, ZNN)。ZNN 是解决时变问题的一种系统方法,在它的设计过程中,与 GNN 相比,ZNN 采用了不同的设计公式、误差函数和动力学方程。在它的实现过程中,ZNN 充分利用了时变系数的导数信息,能够准确有

① 国家自然科学基金(61803338)资助项目。

② 男,1994年生,硕士生;研究方向:机器人轨迹优化;E-mail: gaochangtmzc@163.com。

③ 通信作者,E-mail: kongying-888@163.com。

(收稿日期:2021-03-09)

效地求解时变问题。另外,作为一种新型的递归神经网络,它成功地解决了传统的递归神经网络在求解时变西尔维斯特方程时普遍存在的估计误差问题。根据文献[11],与传统的梯度神经网络相比,ZNN 的优势是随着时间的推移,其估计误差会逐渐减小到 0。

ZNN 一般用于解决实数域中的时变问题,对复数域中的时变问题研究较少。许多时变问题需要考虑到复数,如信号处理领域中频域的在线识别问题^[12]和包含幅值相位信息的输入信号^[13]。为解决上述问题,提出了用于求解复数问题的递归神经网络 (complex-valued zeroing neural network, CVZNN)^[14-16]。这种实时求解复数域时变问题的神经动力学方法可以认为是实值 ZNN 的扩展,它们不仅可以作用于实数域,也可以作用于复数域。但现有的用于求解复数问题的神经网络收敛速度较慢、计算精度较低。根据文献[14],神经网络的收敛情况与初始状态、激励函数的选取和设计参数有关。本文考虑寻找合适的激励函数来提高神经网络的收敛速度和计算精度。

基于以上分析,本文为求解复数域中时变西尔维斯特方程,选取了一种非线性激励函数。根据非线性激励函数在复数域中 2 种等价的处理方法,构建出了 2 种有限时间神经网络 (finite-time neural network, FTNN) 模型。将有限值激励函数应用到 2 种 FTNN 模型中,进一步提高了 2 种 FTNN 模型的收敛速度和计算精度。在求解时变复数西尔维斯特方程时,采用有限值激励函数的 2 种 FTNN 模型收敛速度更快、计算精度更高。并通过稳定性与收敛性的证明验证了 2 种 FTNN 模型的有效性。最后,通过 2 个仿真实例来验证理论推导。

本文的创新点如下。

(1) 在 ZNN 网络模型的基础上,通过复数域中 2 种处理非线性激励函数的方法,构建出 2 种 FTNN 模型,并用它们来求解时变复数西尔维斯特方程。

(2) 与已有的周期神经网络模型相比,FTNN 模型具有有限时间收敛的特性。用它们求解时变复数西尔维斯特方程时效率更高。

(3) 探索了一种有限值激励函数并应用到 2 种

FTNN 模型上,提高了 FTNN 模型的收敛性和稳定性。采用有限值激励函数更易于在实际应用中实现。

1 问题描述和 FTNN 模型

本节首先阐述了一个时变复数西尔维斯特方程的求解问题,然后基于 ZNN 模型在实数域中的设计方法^[14,17-18],再根据非线性激励函数在复数域中 2 种等价的处理方法,构建出了 2 种 FTNN 模型。最后引入了有限值激励函数来提高 2 种 FTNN 模型的收敛速度和计算精度。

1.1 问题描述

首先,给出光滑的时变复数矩阵 $A(t) \in C^{n \times n}$ 、 $B(t) \in C^{n \times n}$ 和 $C(t) \in C^{n \times n}$,求解出满足下列方程的未知复数矩阵 $X(t) \in C^{n \times n}$,方程如下:

$$A(t)X(t) - X(t)B(t) + C(t) = 0 \quad (1)$$

其中 t 表示时间。与普通矩阵不同,矩阵 $A(t)$ 、 $B(t)$ 、 $C(t)$ 和 $X(t)$ 中的每个元素都是复数,并且会随时间的变化而变化,因此每个元素的实部与虚部都会随时间变化。为实时求解出未知的复数矩阵 $X(t)$,需要任意时刻 t 都满足式(1)。神经网络的输入是时变复数矩阵 $A(t)$ 、 $B(t)$ 、 $C(t)$,经过一段时间的演化后,网络中神经元的状态最终会收敛到时变复数西尔维斯特的理论解 $X^*(t)$ 。

基于文献[15]中给出的求解复数域中时变问题的过程,可以将 2 种 FTNN 模型的设计分为以下 3 个步骤。为方便阐述,本文将所构建的 2 种 FTNN 模型分别命名为 FTNN-I 模型和 FTNN-II 模型。

1.2 FTNN-I 模型

矩阵 $A(t)$ 、 $B(t)$ 、 $C(t)$ 和 $X(t)$ 都是由复数构成,每个复数都有各自的实部与虚部。根据第一种处理复数域中非线性激励函数的方法,同时激励复数输入的实部与虚部,按以下步骤构建出 FTNN-I 模型。

步骤 1 为实时监测实际解 $X(t)$ 与理论解 $X^*(t)$ 之间的计算误差,引入动态误差矩阵 $E(t)$ 。随着时间的变化,当 $E(t)$ 趋向于 0 时,此时方程中的 $X(t)$ 就是方程的解, $E(t)$ 如下:

$$\mathbf{E}(t) = \mathbf{A}(t)\mathbf{X}(t) - \mathbf{X}(t)\mathbf{B}(t) + \mathbf{C}(t) \quad (2)$$

步骤 2 为保证动态误差矩阵 $\mathbf{E}(t)$ 最终能收敛到 0, 定义 $\mathbf{E}(t)$ 的动态变化公式为

$$\frac{d\mathbf{E}(t)}{dt} = -\eta(F(\mathbf{E}_{re}(t)) + iF(\mathbf{E}_{im}(t))) \quad (3)$$

其中 $\eta > 0$, $F(\cdot) : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}$ 表示实数域中定义的非线性激励函数, $i = \sqrt{-1}$ 表示虚数单位, $\mathbf{E}_{re}(t) \in \mathbb{R}^{n \times n}$ 和 $\mathbf{E}_{im}(t) \in \mathbb{R}^{n \times n}$ 分别表示误差矩阵 $\mathbf{E}(t) \in \mathbb{C}^{n \times n}$ 的实部与虚部。

步骤 3 联立式(2)和 $\mathbf{E}(t)$ 的动态变化式(3), 得到 FTNN-I 模型:

$$\begin{aligned} \mathbf{A}(t)\dot{\mathbf{X}}(t) - \dot{\mathbf{X}}(t)\mathbf{B}(t) = & -\dot{\mathbf{A}}(t)\mathbf{X}(t) + \mathbf{X}(t)\dot{\mathbf{B}}(t) \\ & \dot{\mathbf{C}}(t) - \eta(F(\mathbf{E}_{re}(t)) + iF(\mathbf{E}_{im}(t))) \end{aligned} \quad (4)$$

其中 $\mathbf{A}(t)$ 、 $\mathbf{B}(t)$ 、 $\mathbf{C}(t)$ 和 $\mathbf{X}(t)$ 的一阶导数分别为 $\dot{\mathbf{A}}(t) = d\mathbf{A}(t)/dt$ 、 $\dot{\mathbf{B}}(t) = d\mathbf{B}(t)/dt$ 、 $\dot{\mathbf{C}}(t) = d\mathbf{C}(t)/dt$ 和 $\dot{\mathbf{X}}(t) = d\mathbf{X}(t)/dt$ 。

1.3 FTNN-II 模型

矩阵 $\mathbf{A}(t)$ 、 $\mathbf{B}(t)$ 、 $\mathbf{C}(t)$ 和 $\mathbf{X}(t)$ 都是由复数构成, 每个复数都有各自的模量与辐角。根据复数域中第二种处理非线性激励函数的方法, 激励复数输入的模量, 按以下步骤构建出 FTNN-II 模型。

步骤 1 与 FTNN-I 模型式(4)的设计过程类似, 引入动态误差矩阵 $\mathbf{E}(t)$:

$$\mathbf{E}(t) = \mathbf{A}(t)\mathbf{X}(t) - \mathbf{X}(t)\mathbf{B}(t) + \mathbf{C}(t) \quad (5)$$

步骤 2 为保证动态误差矩阵 $\mathbf{E}(t)$ 最终能收敛到 0, 定义 $\mathbf{E}(t)$ 的动态变化公式为

$$\frac{d\mathbf{E}(t)}{dt} = -\eta(F(\Theta(t)) \circ \exp(i\Delta(t))) \quad (6)$$

其中 η 、 i 和 $F(\cdot)$ 与上述定义相同, \circ 表示矩阵之间的 Hadamard 乘积, $\Theta(t) \in \mathbb{R}^{n \times n}$ 和 $\Delta(t) \in (-\pi, \pi]^{n \times n}$ 分别表示误差矩阵 $\mathbf{E}(t)$ 的模量与辐角。

步骤 3 联立式(5)和 $\mathbf{E}(t)$ 的动态变化式(6), 得到 FTNN-II 模型:

$$\begin{aligned} \mathbf{A}(t)\dot{\mathbf{X}}(t) - \dot{\mathbf{X}}(t)\mathbf{B}(t) = & -\dot{\mathbf{A}}(t)\mathbf{X}(t) + \mathbf{X}(t)\dot{\mathbf{B}}(t) \\ & -\dot{\mathbf{C}}(t) - \eta(F(\Theta(t)) \circ \exp(i\Delta(t))) \end{aligned} \quad (7)$$

其中, $\dot{\mathbf{A}}(t)$ 、 $\dot{\mathbf{B}}(t)$ 、 $\dot{\mathbf{C}}(t)$ 和 $\dot{\mathbf{X}}(t)$ 分别为 $\mathbf{A}(t)$ 、 $\mathbf{B}(t)$ 、 $\mathbf{C}(t)$ 和 $\mathbf{X}(t)$ 的一阶导数。

从 FTNN-I 模型式(4)和 FTNN-II 模型式(7)的设计过程中可知, FTNN 是一种与 ZNN 相似的递归神经网络。式(3)和式(6)分别描述了 FTNN-I 模型

式(4)和 FTNN-II 模型式(7)中误差矩阵的动态变化过程。从上述 2 个公式中可以看出, 每个时间点估计的误差会实时反馈给神经元, 从而最终使动态误差矩阵 $\mathbf{E}(t)$ 收敛到 0。由于这种实时反馈的特性, 通过 2 种 FTNN 模型求解时变复数西尔维斯特方程问题时, 效率很高。

注 1 复数的模量与辐角, 如一个复数为 $a + bi$, 那么它的模量为 $\sqrt{a^2 + b^2}$, 辐角为 $\arctan \frac{b}{a}$, 辐角有无数个, 但一般选用辐角主值(在区间 $[-\pi, \pi]$ 内的辐角)。

1.4 有限值激励函数

根据文献[14], 求解复数域中的时变问题时, 使用非线性激励函数的 CVZNN 模型可保证误差能以指数级的速度收敛。此外, 根据文献[18], 使用符号双幂激励函数可以提高 CVZNN 模型的收敛速度, 并且能使其在有限时间内收敛。这表明特殊的非线性激励函数可以减少 CVZNN 模型的收敛时间。为进一步提高 2 种 FTNN 模型的收敛速度与计算精度, 尝试了一种非线性激励函数(简称有限值激励函数) $F(\cdot)$:

$$F(x) = x^{\frac{q}{p}}, \quad q \geq 1, p \geq 3, 0 < \frac{q}{p} < 1 \quad (8)$$

其中 q 和 p 都为奇数。有限值激励函数在线性激励函数的基础上, 加了一个指数。根据数学定义, 有限值激励函数仍然是单调递增的奇函数, 所以将它应用到 2 种 FTNN 模型时仍然可以保持全局收敛^[18]。

2 算法的收敛性证明

根据文献[15], 如果使用单调递增的奇函数作为激励函数, 使用 CVZNN 模型来求解时变复数西尔维斯特方程时, 能收敛到理论解。因此, 首先要证明应用有限值激励函数的 2 种 FTNN 模型在求解时变复数西尔维斯特方程时, 能收敛到理论解。然后, 证明其能在有限时间收敛。最后, 计算出收敛时间上界。

2.1 稳定性分析

对于 FTNN-I 模型式(4), 有以下定理。

定理 1 给定光滑的时变复数矩阵 $\mathbf{A}(t)$ 、 $\mathbf{B}(t)$

和 $C(t)$, 如果将有限值激励函数式(8)应用到 FTNN-I 模型式(4)中, 通过 FTNN-I 模型式(4)求解的未知矩阵 $X(t)$ 从任意初始状态 $X(0)$ 开始, 最终能收敛到时变复数西尔维斯特方程的理论解 $X^*(t)$ 。

证明 根据 FTNN-I 模型式(4), 得到其等价形式:

$$\dot{E}(t) = -\eta(F(E_{re}(t)) + iF(E_{im}(t))) \quad (9)$$

其中 $E(t) = A(t)X(t) - X(t)B(t) + C(t)$, $\dot{E}(t)$ 是 $E(t)$ 的一阶导数, $E_{re}(t) \in R^{n \times n}$ 和 $E_{im}(t) \in R^{n \times n}$ 分别表示复数误差矩阵 $E(t)$ 的实部与虚部。将 $\dot{E}(t)$ 的实部和虚部分开:

$$\dot{E}_{re}(t) = -\eta F(E_{re}(t)) \quad (10)$$

$$\dot{E}_{im}(t) = -\eta F(E_{im}(t)) \quad (11)$$

取实部与虚部中的各个元素, 得到另一种形式:

$$\dot{u}_{ij}(t) = -\eta f(u_{ij}(t)) \quad (12)$$

$$\dot{v}_{ij}(t) = -\eta f(v_{ij}(t)) \quad (13)$$

其中, $u_{ij}(t)$ 和 $v_{ij}(t)$ 分别表示 $E_{re}(t)$ 和 $E_{im}(t)$ 的第 i 行第 j 列元素, $\dot{u}_{ij}(t)$ 与 $\dot{v}_{ij}(t)$ 分别表示 $u_{ij}(t)$ 和 $v_{ij}(t)$ 的一阶导数, $f(\cdot)$ 是处理矩阵各个元素的有限值激励函数。定义李亚普诺夫函数 $L_u = \frac{1}{2}u_{ij}^2(t)$, 对其求导可得 $\dot{L}_u = -\eta u_{ij}(t)f(u_{ij}(t))$ 。因为 $f(\cdot)$ 是单调递增的奇函数, 所以有:

$$\dot{L}_u \begin{cases} < 0 & u_{ij} \neq 0 \\ = 0 & u_{ij} = 0 \end{cases} \quad (14)$$

因此, L_u 恒小于等于 0。根据文献[18]中的李亚普诺夫定理, 可以得出, 对于任意的 i, j , $u_{ij}(t)$ 都能全局收敛到 0。对于误差矩阵 $E(t)$ 的虚部, 与上面的证明类似, 也可以证明出 $E_{im}(t)$ 中的任意元素 $v_{ij}(t)$ 最终也都能收敛到 0。也就是说, 误差矩阵 $E(t)$ 随着时间的增加最后都能全局收敛到 0。根据误差矩阵 $E(t)$ 的定义, 当 $E(t)$ 全局收敛到 0 时, 通过 FTNN-I 模型式(4)求出的未知矩阵 $X(t)$ 与初始状态 $X(0)$ 无关, 并最终能收敛到时变复数西尔维斯特方程的理论解 $X^*(t)$ 。

证明成立。

定理 2 给定光滑的时变复数矩阵 $A(t)$ 、 $B(t)$ 和 $C(t)$, 如果将有限值激励函数式(8)应用到 FTNN-II 模型式(7)中, 通过 FTNN-II 模型式(7)求解的未知矩阵 $X(t)$ 从任意初始状态 $X(0)$ 开始, 最终能

收敛到时变复数西尔维斯特方程的理论解 $X^*(t)$ 。

证明 根据 FTNN-II 模型式(7), 得到其等价形式:

$$\dot{E}(t) = -\eta(F(\Theta(t)) \circ \exp(i\Delta(t))) \quad (15)$$

取 $\dot{E}(t)$ 中的各个元素, 可得:

$$\dot{e}_{ij}(t) = -\eta(f(w_{ij}(t)) \exp(i\theta_{ij}(t))) \quad (16)$$

其中, $e_{ij}(t)$ 是复数误差矩阵 $E(t)$ 的第 i 行第 j 列元素, 且 $e_{ij}(t) = w_{ij}(t) \exp(i\theta_{ij}(t))$ 。 $\dot{e}_{ij}(t)$ 是它的一阶导数, $w_{ij}(t)$ 是 $e_{ij}(t)$ 的模数, $\theta_{ij}(t)$ 是 $e_{ij}(t)$ 的辐角。

定义李亚普诺夫函数 $L_e = \frac{1}{2}|e_{ij}(t)|^2 = \frac{1}{2}w_{ij}^2(t)$, 因

为 $e_{ij}(t)$ 是复数, 式子也可以写为 $L_e = \frac{1}{2}|e_{ij}(t)|^2 =$

$\frac{1}{2}e_{ij}(t)\overline{e_{ij}(t)}$, 其中 $\overline{e_{ij}(t)}$ 为 $e_{ij}(t)$ 的共轭复数, 且

$\overline{\dot{e}_{ij}(t)} = \dot{w}_{ij}(t) \exp(-i\theta_{ij}(t))$ 。最后, 对 L_e 求导:

$$\begin{aligned} \dot{L}_e &= \frac{1}{2}(\dot{e}_{ij}(t)\overline{e_{ij}(t)} + e_{ij}(t)\overline{\dot{e}_{ij}(t)}) \\ &= -\frac{1}{2}\eta f(w_{ij}(t)) \exp(i\theta_{ij}(t)) \\ &\quad \cdot w_{ij}(t) \exp(-i\theta_{ij}(t)) \\ &\quad - \frac{1}{2}\eta w_{ij}(t) \exp(i\theta_{ij}(t)) \\ &\quad \cdot f(w_{ij}(t)) \exp(-i\theta_{ij}(t)) \\ &= -\eta w_{ij}(t) f(w_{ij}(t)) \\ &= -\eta |e_{ij}(t)| f(|e_{ij}(t)|) \end{aligned} \quad (17)$$

其中 $\overline{\dot{e}_{ij}(t)} = \dot{f}(w_{ij}(t)) \exp(-i\theta_{ij}(t))$, $\overline{\dot{e}_{ij}(t)}$ 是 $\dot{e}_{ij}(t)$ 的共轭复数。因为 $f(\cdot)$ 是单调递增的奇函数, 所以可以推出 L_e 恒小于等于 0。根据李亚普诺夫定理, 可以得出对于任意的 i, j , $e_{ij}(t)$ 最终都能全局收敛到 0。根据误差矩阵 $E(t)$ 的定义, 当 $E(t)$ 趋于 0 时, 通过 FTNN-II 模型(7)求出的未知矩阵 $X(t)$ 与初始状态 $X(0)$ 无关, 并最终能收敛到时变复数西尔维斯特方程的理论解 $X^*(t)$ 。

证明成立。

注 2 从定理 1 和定理 2 的证明过程来看, 下面的这些激励函数同样满足 $F(\cdot)$ 的要求(单调递增的奇函数), 并可以应用到 FTNN-I 模型式(4)和 FTNN-II 模型式(7)上。

线性激励函数:

$$F(x) = x \quad (18)$$

幂函数:

$$F(x) = x^k, k \geq 3 \quad (19)$$

双 S 型函数:

$$F(x) = \frac{1 + \exp(-m)}{1 - \exp(-m)} \frac{1 - \exp(-mx)}{1 + \exp(-mx)}, m > 2 \quad (20)$$

幂-S 型函数:

$$F(x) = \frac{1}{2}x^k + \frac{1}{2} \frac{1 + \exp(-m)}{1 - \exp(-m)} \frac{1 - \exp(-mx)}{1 + \exp(-mx)}, k \geq 3, m > 2 \quad (21)$$

符号双幂函数:

$$F(x) = \frac{1}{2} |x|^r \text{sign}(x) + \frac{1}{2} |x|^{\frac{1}{r}} \text{sign}(x), 0 < r < 1 \quad (22)$$

在下文的仿真部分,将这些激励函数与有限值激励函数式(8)应用到 2 种 FTNN 模型中,再分别进行比较。

2.2 收敛性分析

根据文献[18],在实数域内,如果使用符号双幂激励函数,那么就能在有限时间内求解出时变西尔维斯特方程,并且可以计算出理论解的收敛时间上界。本节将有限值激励函数从实数域扩展到了复数域,以便能在有限时间内求解出时变复数西尔维斯特方程,并计算出理论解的收敛时间上界。

定理 3 给定光滑的时变矩阵 $\mathbf{A}(t)$ 、 $\mathbf{B}(t)$ 和 $\mathbf{C}(t)$,如果将有限值激励函数式(8)应用到 2 种 FTNN 模型上,则通过 2 种 FTNN 模型求解出的未知矩阵 $\mathbf{X}(t)$ 无论从任意初始状态 $\mathbf{X}(0)$ 开始,最终能在一段时间后收敛到时变复数西尔维斯特方程的理论解 $\mathbf{X}^*(t)$ 。FTNN-I 模型式(4)和 FTNN-II 模型式(7)的收敛时间上界分别如下所示。

对于 FTNN-I 模型式(4),它的收敛时间上界 t_1 为

$$t_1 = \frac{p}{\eta(p-q)} m(0)^{\frac{p-q}{p}} \quad (23)$$

其中, $m(0) = \max\{|u_{ij}(0)|, |v_{ij}(0)|\}$, $u_{ij}(0)$ 和 $v_{ij}(0)$ 分别表示误差矩阵的初始状态 $\mathbf{E}(0)$ 中第 i 行第 j 列元素的实部与虚部。

对于 FTNN-II 模型式(7),它的收敛时间上界 t_2 为

$$t_2 = \frac{p}{\eta(p-q)} n(0)^{\frac{p-q}{p}} \quad (24)$$

其中, $n(0) = \max\{|e_{ij}(0)|\}$, $e_{ij}(0)$ 表示误差矩阵的初始状态 $\mathbf{E}(0)$ 的第 i 行第 j 列元素。

证明 考虑到 2 种 FTNN 模型,证明过程分以下 2 种情况。

情况 1 对于 FTNN-I 模型式(4),根据式(9),可得:

$$\dot{\mathbf{E}}(t) = -\eta(F(\mathbf{E}_{re}(t)) + iF(\mathbf{E}_{im}(t))) \quad (25)$$

再根据式(12)和式(13),可得:

$$\dot{u}_{ij}(t) = -\eta f(u_{ij}(t)) \quad (26)$$

$$\dot{v}_{ij}(t) = -\eta f(v_{ij}(t)) \quad (27)$$

在初始时间中选取 $u_{ij}(t)$ 和 $v_{ij}(t)$ 中较大的一个作为 $m(t)$ 的初值, $m(0) = \max\{|u_{ij}(0)|, |v_{ij}(0)|\}$ 。在任意时刻 t 中,取 $u_{ij}(t)$ 和 $v_{ij}(t)$ 中绝对值相对较大的一个作为 $m(t)$ 。通过比较定理,可以得到:

$-|m(t)| \leq u_{ij}(t) \leq |m(t)|$, $-|m(t)| \leq v_{ij}(t) \leq |m(t)|$ 。换句话说,当 $m(t)$ 等于 0 时, $\mathbf{E}_{re}(t)$ 和 $\mathbf{E}_{im}(t)$ 中所有的元素都会等于 0。因为 $f(\cdot)$ 是一个单调递增的奇函数, $\dot{\mathbf{E}}_{re}(t)$ 中所有的元素 $u_{ij}(t)$ 都满足式(26), $\dot{\mathbf{E}}_{im}(t)$ 中所有的元素 $v_{ij}(t)$ 都满足式(27),所以可以得到 $m(t)$ 的动态表达式为

$$\dot{m}(t) = -\eta f(m(t)) \quad (28)$$

定义李亚普诺夫函数 $L(t) = |m(t)|^2$,对它求导,得到:

$$\begin{aligned} \dot{L}(t) &= -2\eta m(t)f(m(t)) = -2\eta \overline{m}(t)(m^{\frac{a}{p}}(t)) \\ &= -2\eta m^{\frac{q+p}{p}}(t) = -2\eta L^{\frac{q+p}{2p}}(t) \end{aligned} \quad (29)$$

解上述微分方程式(29),得到:

$$\frac{d(L(t))}{L^{\frac{q+p}{2p}}(t)} = -2\eta dt \quad (30)$$

将微分方程式(30)从 0 积分到 t ,可得:

$$L(t) = \left(-\frac{\eta(p-q)}{p}t + L(0)^{\frac{p-q}{2p}}\right)^{\frac{2p}{p-q}} \quad (31)$$

因为 $L(t) = |m(t)|^2$,所以 $L(t) \geq 0$,故:

$$\left(-\frac{\eta(p-q)}{p}t + L(0)^{\frac{p-q}{2p}}\right) \geq 0 \quad (32)$$

因此可以解出:

$$t_1 \leq \frac{p}{\eta(p-q)} m(0)^{\frac{p-q}{p}} \quad (33)$$

所以收敛时间上界 $t_1 = \frac{p}{\eta(p-q)} m(0)^{\frac{p-q}{p}}$ 。

情况 2 定义李亚普诺夫函数 $L(t) = |n(t)|^2$, 其中 $n(0)$ 取 $\mathbf{E}(t)$ 所有元素中模数的最大值, 且初始状态 $n(0) = \max\{|e_{ij}(0)|\}$ 。可以得出结论, 当 $n(t)$ 趋于 0 时, $\mathbf{E}(t)$ 中所有的元素都会等于 0。因此, 如果计算出了 $n(t)$ 的收敛时间, 就能相应地估算出 FTNN-II 模型式 (7) 的收敛时间上界。根据式 (6), 取矩阵中的各个元素, 可以得到 $n(t)$ 的动态表达式为

$$\dot{n}(t) = -\eta(f(w_{ij}(t)) \exp(i\theta_{ij}(t))) \quad (34)$$

类似式 (17) 的推导, 可以得出:

$$\begin{aligned} \dot{L}(t) &= -2\eta n(t)f(n(t)) = -2\eta n(t)(n^{\frac{q}{p}}(t)) \\ &= -2\eta n^{\frac{q+p}{p}}(t) = -2\eta L^{\frac{q+p}{2p}}(t) \end{aligned} \quad (35)$$

与情况 1 的证明类似, FTNN-II 模型式 (7) 的收敛时间上界计算结果为 $t_2 = \frac{p}{\eta(p-q)}n(0)^{\frac{p-q}{p}}$ 。

证明成立。

3 数值仿真

本节给出了 2 个具体实例, 首先将有限值激励函数式 (8) 应用到 FTNN-I 模型式 (4) 和 FTNN-II 模型式 (7) 中; 然后将其他单调递增的奇函数与有限值激励函数式 (8) 进行比较; 最后为证明有限值激励函数式 (8) 能减少 FTNN-I 模型式 (4) 和 FTNN-II 模型式 (7) 的收敛时间, 本文单独将有限值激励函数式 (8) 与符号双幂激励函数式 (22) 进行比较。

仿真 1 求解时变复数西尔维斯特方程: $\mathbf{A}(t)\mathbf{X}(t) - \mathbf{X}(t)\mathbf{B}(t) + \mathbf{C}(t) = \mathbf{0}$, 其中 $\mathbf{A}(t)$ 、 $\mathbf{B}(t)$ 和 $\mathbf{C}(t)$ 的系数如下, 求未知矩阵 $\mathbf{X}(t)$ 。

$$\mathbf{A}(t) = \begin{bmatrix} \exp(i5t) & -\cos(5t) + \sin(5t)i \\ -\exp(i5t) & 0 \end{bmatrix},$$

$$\mathbf{B}(t) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \mathbf{C}(t) = -\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}。$$

此时, $\mathbf{X}(t)$ 的逻辑解为

$$\mathbf{X}^*(t) = -\begin{bmatrix} 0 & \exp(-i5t) \\ \cos(5t) + \sin(5t)i & \cos(5t) + \sin(5t)i \end{bmatrix}。$$

图 1 描述了当 FTNN-I 模型式 (4) 使用有限值激励函数式 (8), 并用求解时变西尔维斯特方程时, 未知矩阵 $\mathbf{X}(t)$ 的阶段性变化 (FTNN-I 模型

式 (4) 中设定 $\eta = 50, \frac{q}{p} = \frac{1}{3}$)。其中实线表示实际解的状态变化, 虚线表示理论解 $\mathbf{X}^*(t)$ 的状态变化。图中实线与虚线会在一段时间后重合, 这说明 $\mathbf{X}(t)$ 从初始状态 $\mathbf{X}(0)$ 开始 (仿真中初步设定初值 $\mathbf{X}(0) = [1+i, 1+i; 0, 0]$, 最终能在一段时间后收敛到时变复数西尔维斯特方程的理论解。

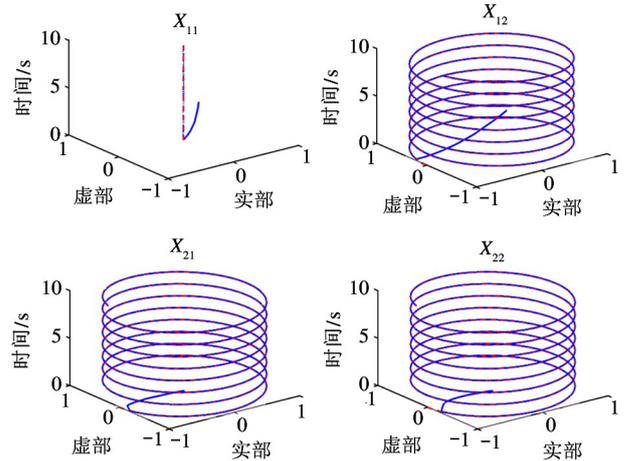


图 1 $\mathbf{X}(t)$ 的理论解 (虚线) 和实际解 (实线)

一般情况下, 实际解 $\mathbf{X}(t)$ 与理论解 $\mathbf{X}^*(t)$ 之间存在计算误差, 用 $\|\mathbf{A}(t)\mathbf{X}(t) - \mathbf{X}(t)\mathbf{B}(t) + \mathbf{C}(t)\|_F$ 来近似估计实际解与理论解之间的计算误差。然后, 将线性激励函数、幂激励函数 ($k=3$)、双 S 型激励函数 ($m=3$)、幂-S 型激励函数 ($k=3, m=3$)、符号双幂激励函数 ($r=3$) 与有限值激励函数 ($\frac{q}{p} =$

$\frac{1}{3}$) 进行比较。为进行对比验证, 将 FTNN-I 模型式 (4) 中的设计参数 η 统一设置为 50。如图 2 所示, 当这些激励函数应用在 FTNN-I 模型式 (4) 时都能使计算误差逐渐减小。这验证了前面的理论推导: 即如果激励函数是单调递增的奇函数, 那么就能保证全局收敛。此外, 从图中可以看出, 采用有限值激励函数式 (8) 的 FTNN-I 式 (4) 模型的误差收敛速度比其他激励函数要快, 这说明在求解时变复数西尔维斯特方程时, 有限值激励函数的效果要更好。

因为符号双幂激励函数式 (22) 与有限值激励函数式 (8) 是指数级的函数, 误差的收敛速度较快, 并且容易计算出收敛时间上界, 所以在 FTNN-I 模型

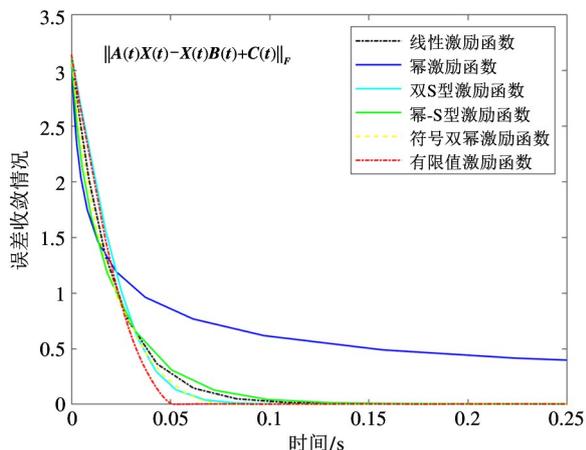


图2 激励函数的收敛性比较

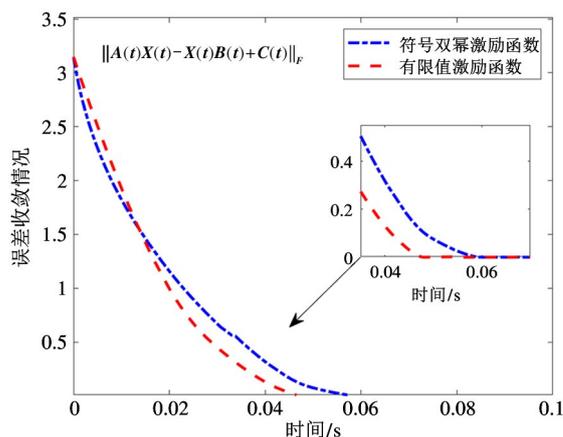


图3 符号双幂激励函数与有限值激励函数的收敛性比较

式(4)中,单独将符号双幂激励函数式(22)与有限值激励函数式(8)进行比较。为计算出收敛时间上界,使用的是特定的初始状态而不是随机产生的初始状态。设定 $X(0) = [1 + i, 1 + i; 0, 0]$, 误差矩阵的初始状态 $E(0) = A(0)X(0) - X(0)B(0) + C(0)$, 计算为 $E(0) = [i, 1 + i; -1 - i, -2 - i]$, 则式(23)中定义的最大项 $m(0) = 3$ 。将有限值激励函数式(8) (设定 $\eta = 50, \frac{q}{p} = \frac{1}{3}$) 应用到 FTNN-I 模型式(4)中, 根据式(23), 计算出使用有限值激励函数的 FTNN-I 模型式(4)的收敛时间上界 $t_1 = \frac{p}{\eta(p-q)} m(0)^{\frac{p-q}{p}} \approx 0.0476 \text{ s}$ 。再将符号双幂激励函数式(22) (设定 $\eta = 50, \frac{q}{p} = \frac{1}{3}$) 应用到到 FTNN-I 模型式(4)中, 根据文献[14]中的定理3, 计算出其收敛时间上界 $t_2 = \frac{2|c(0)|^{1-r}}{\eta(1-r)} \approx 0.0928 \text{ s}$ (r 取0.3时, 符号双幂激励函数的收敛情况较好)。

如图3所示,在 FTNN-I 式(4)模型中,使用符号双幂激励函数式(22)的理论收敛时间上界为 0.0928 s, 而仿真实际运行时间为 0.0586 s。使用有限值激励函数式(8)的理论收敛时间上界为 0.0476 s, 而仿真实际运行时间为 0.0467 s, 符合理论分析。当 FTNN-I 模型式(4)使用有限值激励函数式(8)时, 误差收敛速度要比使用符号双幂激励函数式(22)时更快。这也证明了在 FTNN-I 模型式(4)中, 用有限值激励函数式(8)来求解时变复数西尔维斯特方程时, 效果更好。

接下来讨论设计参数 η 对模型的影响, 先固定 $\frac{q}{p} = \frac{1}{3}$, 再判断不同的 η 对收敛性能的影响。

图4描述了用 FTNN-I 模型式(4)来求解时变复数西尔维斯特方程时, 设计参数 η 对误差收敛速度的影响。当增加 η 的值时, 可以提高模型误差的收敛速度。

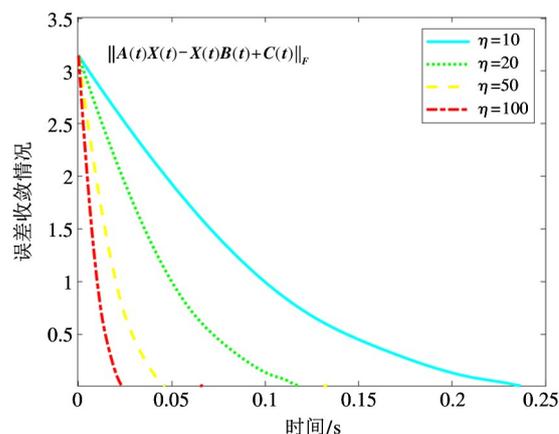


图4 设计参数 η 的收敛性比较

仿真2 为进一步验证理论的有效性, 举第2个仿真实例。求解时变复数西尔维斯特方程: $A(t)X(t) - X(t)B(t) + C(t) = 0$, 其中 $A(t)$ 、 $B(t)$ 和 $C(t)$ 的系数如下, 求未知矩阵 $X(t)$ 。

$$A(t) = \begin{bmatrix} \exp(-it) & -i\exp(-it) \\ -\exp(-it) & \exp(-it) \end{bmatrix},$$

$$B(t) = \begin{bmatrix} -2i\exp(2it) & \exp(-it) \\ \exp(it) & -2i\exp(-2it) \end{bmatrix},$$

$$C(t) = - \begin{bmatrix} -i\exp(it) - \frac{1}{2} & \frac{\exp(-2it)}{2} + \exp(-3it) \\ \frac{\exp(2it)}{2} + i\exp(3it) + 1 & -i\exp(-it) - 1 + \frac{i}{2} \end{bmatrix}$$

此时, $X(t)$ 的理论解为

$$X^*(t) = - \begin{bmatrix} \frac{1}{2}\exp(-it) & \frac{1}{2}i\exp(-it) \\ -\frac{1}{2}\exp(it) & \frac{1}{2}\exp(it) \end{bmatrix}$$

类似于仿真 1, 图 5 描述了当 FTNN-II 模型式(7)使用有限值激励函数式(8), 并用来求解时变西尔维斯特方程时, 未知矩阵 $X(t)$ 的阶段性变化 (FTNN-II 模型式(7)中设定 $\eta = 50, \frac{q}{p} = \frac{1}{3}$)。图中实线(实际解)与虚线(理论解)会在一段时间后重合。这说明 $X(t)$ 从初始状态 $X(0)$ 开始(仿真中初步设定初值 $X(0) = [0.5i, 0; 0, 0]$, 最终能收敛到时变复数西尔维斯特方程的理论解。

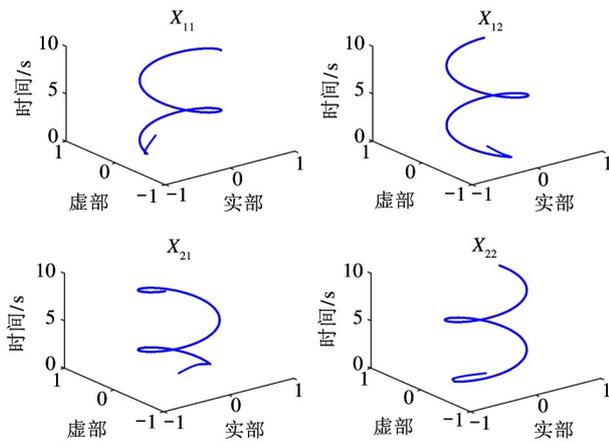


图 5 $X(t)$ 的理论解(虚线)和实际解(实线)

图 6 描述了在 FTNN-II 模型式(10)中, 使用一些激励函数与有限值激励函数式(8)进行收敛性比较, 其中这些激励函数的设计参数与仿真一完全相同。从图中可以看出, 使用有限值激励函数式(8)的 FTNN-II 模型式(7)的误差收敛速度比使用其他激励函数时要更快。

在 FTNN-II 模型式(7)中, 单独将符号双幂激励函数式(22)与有限值激励函数式(8)进行误差的收敛性比较。设定初始状态 $X(0) = [0.5i, 0; 0, 0]$, 误差矩阵的初始状态 $E(0) = A(0)X(0) - X(0)$

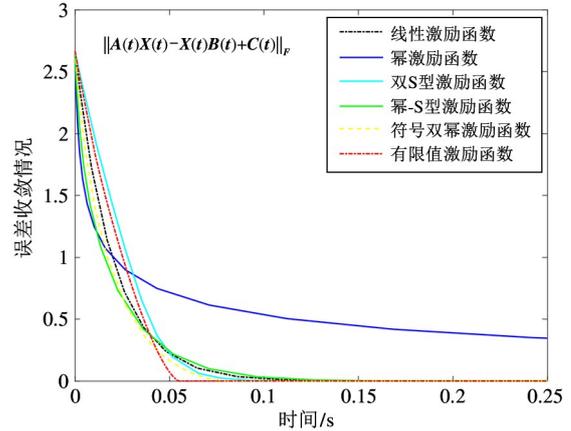


图 6 激励函数的收敛性比较

$B(0) + C(0)$, 计算为: $E(0) = [1.5i - 0.5, -0.5i - 1.5; -1.5i - 1.5, 0.5i + 1]$, 则式(24)中定义的最大项 $n(0) = \frac{3\sqrt{2}}{2}$, 将有限值激励函数式(8) (设定 $\eta = 50, \frac{q}{p} = \frac{1}{3}$) 应用到 FTNN-II 模型式(7)中, 根据式(24), 计算出使用 FTNN-II 模型式(7)的收敛时间上界 $t_1 = \frac{p}{\eta(p-q)} n(0)^{\frac{p-q}{p}} \approx 0.0495$ s。再将符号双幂激励函数式(22) (设定 $\eta = 50, r = 0.3$) 应用到 FTNN-II 模型式(7)中, 通过文献[14]中的定理 3, 计算出收敛时间上界 $t_2 = \frac{2 \|c(0)\|^{1-r}}{\eta(1-r)} \approx 0.0967$ s。

如图 7 所示, 在 FTNN-II 式(7)模型中, 使用符号双幂激励函数式(22)的理论收敛时间上界为 0.0967 s, 而仿真实际运行时间为 0.0577 s。使用有

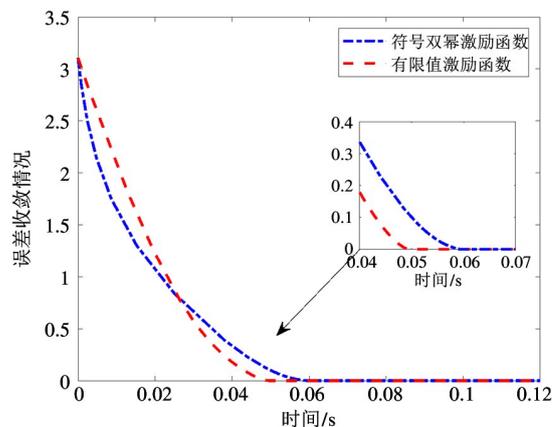


图 7 符号双幂激励函数与有限值激励函数的收敛性比较

限值激励函数式(8)的理论收敛时间上界为 0.0495 s,而仿真实际运行时间为 0.0484 s,符合理论分析。当 FTNN-II 模型式(7)使用有限值激励函数式(8)时,误差收敛速度要比使用符号双幂激励函数式(22)更快。这也证明了在 FTNN-II 模型式(7)中,使用有限值激励函数式(8)时,效果更好。

图 8 描述了用 FTNN-II 模型式(7)来求解时变复数西尔维斯特方程时,设计参数 η 对误差收敛速度的影响。如图所示, η 越大,误差的收敛速度越快。

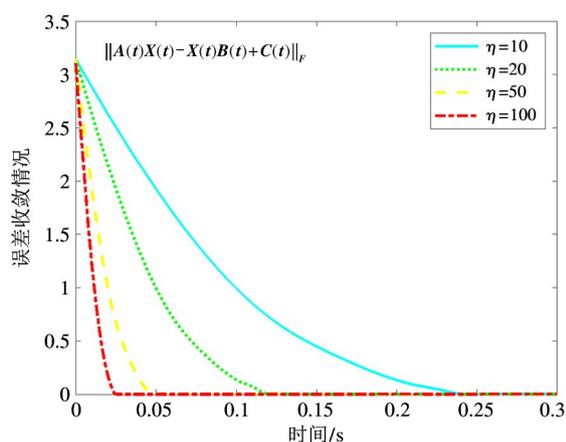


图 8 设计参数 η 的收敛性比较

4 结 论

本文针对时变复数西尔维斯特方程的求解,首先基于复数的运算规则,设计了 2 种改进的有限时间神经网络模型。理论分析表明,当使用任意单调递增的奇函数作为激励函数时,这 2 种 FTNN 模型都可以实现全局收敛性。然后使用有限值激励函数来提高 2 种 FTNN 模型的收敛速度与计算精度。最后将符号双幂激励函数与有限值激励函数应用到 2 种 FTNN 模型中,与符号双幂激励函数相比,应用有限值激励函数的 FTNN 模型收敛时间上界更小。数值算例与仿真结果验证了所提出的 2 种基于有限值激励函数的 FTNN 模型在求解时变复数西尔维斯特方程时的优越性。然而,由于 2 种 FTNN 模型中需要用到非线性激励函数并进行数值运算,因此 FTNN 模型比应用线性激励函数的 CVZNN 模型具有更高的复杂度,未来的工作可能会优化 FTNN 模型

的结构,并进一步推广到实际应用中。

参考文献

- [1] CALVETTI D, REICHEL L. Application of ADI iterative methods to the restoration of noisy images [J]. *SIAM Journal on Matrix Analysis and Applications*, 1996, 17 (1): 165-186
- [2] XIAO L, ZHANG Z, ZHANG Z, et al. Design, verification and robotic application of a novel recurrent neural network for computing dynamic Sylvester equation [J]. *Neural Networks*, 2018, 105: 185-196
- [3] BRAHMA S, DATTA B. An optimization approach for minimum norm and robust partial quadratic eigenvalue assignment problems for vibrating structures [J]. *Journal of Sound and Vibration*, 2009, 324 (3-5): 471-489
- [4] ZHANG Y, WANG J. Global exponential stability of recurrent neural networks for synthesizing linear feedback control systems via pole assignment [J]. *IEEE Transactions on Neural Networks*, 2002, 13 (3): 633-644
- [5] ZHANG Y, JIANG D, WANG J. A recurrent neural network for solving Sylvester equation with time-varying coefficients [J]. *IEEE Transactions on Neural Networks*, 2002, 13 (5): 1053-1063
- [6] TANG C, TAO Y, CHEN Y, et al. Approximate iteration detection and precoding in massive MIMO [J]. *China Communications*, 2018, 15 (5): 183-196
- [7] 孙明轩,吴雨芯,张钰. 二次曲线型终态神经网络:时变神经计算与冗余机械臂重复运动规划 [J]. *高技术通讯*, 2020, 30 (10): 59-70
- [8] 陈劲峰,黄卫华,王肖,等. 基于改进蚁群算法的移动机器人路径规划 [J]. *高技术通讯*, 2020, 30 (3): 291-297
- [9] XIAO L, ZHANG Y. Zhang neural network versus gradient neural network for solving time-varying linear inequalities [J]. *IEEE Transactions on Neural Networks*, 2011, 22 (10): 1676-1684
- [10] ZHANG Y, GE S S. Design and analysis of a general recurrent neural network model for time-varying matrix inversion [J]. *IEEE Transactions on Neural Networks*, 2005, 16 (6): 1477-1490
- [11] ZHANG Y, YI C, GUO D, et al. Comparison on Zhang neural dynamics and gradient-based neural dynamics for online solution of nonlinear time-varying equation [J].

- Neural Computing and Applications*, 2011, 20(1): 1-7
- [12] SUBRAMANIAN K, SAVITHA R, SURESH S. A complex-valued neuro-fuzzy inference system and its learning mechanism[J]. *Neurocomputing*, 2014, 123: 110-120
- [13] DURÁN-DÍAZ I, CRUCES S, SARMIENTO-VEGA M A, et al. Cyclic maximization of non-Gaussianity for blind signal extraction of complex-valued sources [J]. *Neurocomputing*, 2011, 74(17): 2867-2873
- [14] LI S, LI Y. Nonlinearly activated neural network for solving time-varying complex Sylvester equation [J]. *IEEE transactions on Cybernetics*, 2013, 44(8): 1397-1407
- [15] ZHANG Y, LI Z, LI K. Complex-valued Zhang neural network for online complex-valued time-varying matrix inversion [J]. *Applied Mathematics and Computation*, 2011, 217(24): 10066-10073
- [16] QIAO S, WANG X Z, WEI Y. Two finite-time convergent Zhang neural network models for time-varying complex matrix Drazin inverse[J]. *Linear Algebra and its Applications*, 2018, 542: 101-117
- [17] XIAO L, LIAO B, LI S, et al. Nonlinear recurrent neural networks for finite-time solution of general time-varying linear matrix equations[J]. *Neural Networks*, 2018, 98: 102-113
- [18] SHEN Y, MIAO P, HUANG Y, et al. Finite-time stability and its application for solving time-varying Sylvester equation by recurrent neural network[J]. *Neural Processing Letters*, 2015, 42(3): 763-784

Finite-time neural network study for time-varying complex Sylvester equation

GAO Chang, KONG Ying

(School of Information and Electronic Engineering, Zhejiang University of Science and Technology, Hangzhou 310023)

Abstract

In view of solving the real-time problem of the time-varying complex Sylvester equation, two finite-time neural network (FTNN) models are proposed, which is based on the designing rule of Zhang neural network (ZNN) in the real calculation domain. The designing rule of FTNN is aiming at the neural dynamics equation for the complex number domain. Two equivalent methods are applied to the numerical calculation of nonlinear excitation function in dynamic equation. The first method is to deal with the real part and imaginary part of the complex input, and the second method is to deal with the modulus of the complex input. The finite-value activation function is used to accelerate the convergent speed of the FTNN model and obtain high calculation precision. The experimental results show that the proposed network model has better convergence and stability for the solution of time-varying complex Sylvester equation compared with the traditional methodology.

Key words: time-varying complex Sylvester equation, finite-time neural network (FTNN), finite-value activation function, convergent speed, calculation precision