

基于有限时间神经网络求解的时变复数矩阵方程

高 畅,孔 颖,胡汤珑

(浙江科技学院 信息与电子工程学院,杭州 310023)

摘 要: 为求解时变复数矩阵方程,根据复数域中两种等价处理非线性激励函数的方法,提出了两种新型有限时间归零神经网络(new finite-time zeroing neural network,NFTZNN)模型。尝试将一种新型激励函数应用到两种 NFTZNN 模型中,从而提高了模型的综合性能。试验结果表明,与现有的复数神经网络(complex-value zeroing neural network,CVZNN)模型相比,使用新型激励函数的 NFTZNN 模型在求解时变复数矩阵方程时,收敛速度更快、计算精度更高;并且,根据李亚普洛夫定理计算出的收敛时间上界也更接近实际的收敛时间。本研究提出的神经网络模型能准确快速地求解出复数域中的时变矩阵方程,可以为后续工程应用提供参考。

关键词: 矩阵方程;有限时间归零神经网络;新型激励函数;收敛时间上界

中图分类号: TP183 **文献标志码:** A **文章编号:** 1671-8798(2022)05-0409-10

Solving time-varying complex matrix equations based on finite-time neural network

GAO Chang, KONG Ying, HU Tanglong

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

Abstract: Aiming at solving the time-varying complex matrix equations, two new finite-time zeroing neural network (NFTZNN) models were proposed according to two equivalent methods to deal with nonlinear activation functions in the complex domain. A new activation function was applied to two NFTZNN models to improve the comprehensive performance of the models. The experimental results show that compared with the existing complex-value zeroing neural network (CVZNN) model, the NFTZNN models using the new activation function boast faster convergence speed and higher calculation accuracy when solving time-varying complex matrix equations. Moreover, the upper bound of the convergence time calculated according to the

收稿日期: 2021-07-12

基金项目: 国家自然科学基金项目(61803338)

通信作者: 孔 颖(1980—),女,浙江省杭州人,副教授,博士,主要从事机器人轨迹优化研究。E-mail:kongying888@163.com。

Lyapunov theorem is also closer to the actual convergence time. The proposed neural network models can accurately and quickly solve the time-varying matrix equations in the complex field, which can provide important reference for subsequent engineering applications.

Keywords: matrix equation; finite-time zeroing neural network; new activation function; upper bound of convergence time

矩阵方程求解在工程实践中有着非常广泛的应用场景,例如多进多出(multiple-input multiple-output, MIMO)技术^[1]、现场可编程逻辑门阵列(field-programmable gate array, FPGA)设计^[2]和机械臂应用^[3]。矩阵方程的求解主要有两种方法:第一种方法是使用基于迭代思想产生的传统数值算法,如近似迭代法^[4]、蒙特卡罗方法^[5]和插值分解法^[6],由于数值算法计算的串行性,使用它们来处理大规模数据和高维数矩阵时,迭代计算的性能可能不够高;第二种方法是使用一些神经网络算法^[7-8],神经网络算法具有高速并行处理数据的能力且能够实时运算,因此在求解高维数矩阵方程和复杂的时变矩阵方程时,计算性能较高。

基于标量取值能量函数的梯度神经网络(the gradient-based neural network, GNN)是一种传统的神经网络,可用于求解矩阵方程。然而, GNN 在求解过程中基本上都只考虑静态情况,它不会考虑时变情况造成的影响。因此在求解过程中,它的估计误差可能不会收敛到零。因此,使用该方法求解时变矩阵方程时,精度不高。

为实时求解时变矩阵方程,张雨浓提出了一种归零神经网络(zeroing neural network, ZNN)^[9-10]。ZNN 是一种递归神经网络,它可以有效地解决 GNN 中存在的估计误差收敛问题^[11],它的估计误差会随时间的变化而最终收敛到零。因此, ZNN 能得到时变矩阵方程的准确解。但是, ZNN 只能用于求解实数域中的时变矩阵方程,不能作用于复数域。目前,许多时变问题需要进一步考虑到复数域的情况。如计算机视觉中的图像处理^[12]和计算机科学中的模式识别^[9]。针对上述问题,一些研究者提出了能应用于复数域中的神经网络(complex-valued zeroing neural network, CVZNN)^[11, 13-14]。但用 CVZNN 求解时变复数问题时,收敛速度较慢、计算精度不高。根据文献[13],神经网络的收敛速度、计算精度与网络的设计参数和选取的激励函数有关。因此,可以考虑选取实数域中的一些激励函数,将其扩展到复数域,以此来提高神经网络的稳定性和收敛性。

针对时变复数矩阵方程的求解问题,本文首先提出两种新型有限时间归零神经网络(new finite-time zeroing neural network, NFTZNN)模型,该模型在求解上述问题时性能较好。然后,根据李亚普洛夫定理计算出收敛时间上界,再进行稳定性和收敛性的证明来验证理论分析的有效性。最后,通过仿真试验,可验证理论推导的可行性。

1 问题提出和 ZNN 模型

一般情况下,定义一个时变复数矩阵方程为

$$\mathbf{F}(t)\mathbf{Z}(t) = \mathbf{I}. \quad (1)$$

式(1)中: t 表示时间。给定光滑的时变复数矩阵 $\mathbf{F}(t) \in \mathbb{C}^{n \times n}$ 和单位矩阵 $\mathbf{I} \in \mathbb{C}^{n \times n}$,求满足方程的复数矩阵 $\mathbf{Z}(t) \in \mathbb{C}^{n \times n}$ 。上述矩阵的维数都是 n 行 n 列。因为矩阵 $\mathbf{F}(t)$ 和 $\mathbf{Z}(t)$ 是时变复数矩阵,矩阵中的所有元素都是复数且会随时间变化,所以求解出的 $\mathbf{Z}(t)$ 需要在任意时刻 t 都满足式(1)。

为实时求解时变复数矩阵方程(1),仿照实值域中 ZNN 的设计方法^[13, 15],构造出能应用于复数域的标准 ZNN 模型。

首先,为估计实际解 $\mathbf{Z}(t)$ 与理论解 $\mathbf{Z}^*(t)$ 之间的计算误差,定义误差矩阵 $\mathbf{E}(t) \in \mathbb{C}^{n \times n}$ 。当 $\mathbf{E}(t)$ 趋于零时,此时方程中的 $\mathbf{Z}(t)$ 就是方程的理论解 $\mathbf{Z}^*(t)$,误差矩阵

$$\mathbf{E}(t) = \mathbf{F}(t)\mathbf{Z}(t) - \mathbf{I}. \quad (2)$$

然后,引入标准 ZNN 设计公式:

$$\dot{\mathbf{E}}(t) = -\eta \Phi(\mathbf{E}(t)). \quad (3)$$

式(3)中: $\dot{\mathbf{E}}(t)$ 为误差矩阵 $\mathbf{E}(t)$ 的一阶导数;设计参数 $\eta > 0$; $\Phi(\cdot): \mathbb{C}^{n \times n} \rightarrow \mathbb{C}^{n \times n}$ 为非线性激励函数,作用范围是复数域。

最后,联立式(2)和式(3),得到时变复数矩阵方程(1)的标准 ZNN 模型为

$$\dot{\mathbf{E}}(t) = -\eta \Phi(\mathbf{F}(t)\mathbf{Z}(t) - \mathbf{I}). \quad (4)$$

2 NFTZNN 模型

首先在标准 ZNN 模型(4)的基础上,将实数域中的非线性激励函数扩展到复数域,根据复数域中两种等价地处理非线性激励函数的方法,构建出两种 NFTZNN 模型。然后,引入了一种新型激励函数来提高两种 NFTZNN 模型的收敛速度和计算精度。最后,将两种新构建的 NFTZNN 模型分别命名为 NFTZNN-I 模型和 NFTZNN-II 模型。

2.1 NFTZNN-I 模型

处理式(4)中非线性激励函数的第一种方法是同时处理复数输入的实部与虚部,具体定义描述如下:

$$\Phi_1(\mathbf{E}(t)) = \Omega(\mathbf{U}(t)) + i\Omega(\mathbf{V}(t)). \quad (5)$$

式(5)中: $\Omega(\cdot): \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}$ 为非线性激励函数,作用范围是实数域; $i = \sqrt{-1}$ 表示虚数单位; $\mathbf{U}(t) \in \mathbb{R}^{n \times n}$ 和 $\mathbf{V}(t) \in \mathbb{R}^{n \times n}$ 分别是误差矩阵 $\mathbf{E}(t) \in \mathbb{C}^{n \times n}$ 的实部与虚部。

联立式(3)和式(5),可得:

$$\dot{\mathbf{E}}(t) = -\eta(\Omega(\mathbf{U}(t)) + i\Omega(\mathbf{V}(t))). \quad (6)$$

将式(6)中 $\dot{\mathbf{E}}(t)$ 展开,得到 NFTZNN-I 模型:

$$\mathbf{F}(t) \dot{\mathbf{Z}}(t) = -\dot{\mathbf{F}}(t)\mathbf{Z}(t) - \eta(\Omega(\mathbf{U}(t)) + i\Omega(\mathbf{V}(t))). \quad (7)$$

2.2 NFTZNN-II 模型

处理式(4)中非线性激励函数的第二种方法是处理复数输入的模,具体定义描述如下:

$$\Phi_2(\mathbf{E}(t)) = (\Omega(\mathbf{C}(t)) \circ \exp(i\mathbf{D}(t))). \quad (8)$$

式(8)中: \circ 为矩阵之间的哈达玛积; $\mathbf{C}(t)$ 和 $\mathbf{D}(t)$ 分别为误差矩阵 $\mathbf{E}(t)$ 的模与辐角。

联立式(3)和式(8),可得

$$\dot{\mathbf{E}}(t) = -\eta(\Omega(\mathbf{C}(t)) \circ \exp(i\mathbf{D}(t))). \quad (9)$$

将式(9)中 $\dot{\mathbf{E}}(t)$ 展开,得到 NFTZNN-II 模型:

$$\mathbf{F}(t) \dot{\mathbf{Z}}(t) = -\dot{\mathbf{F}}(t)\mathbf{Z}(t) - \eta(\Omega(\mathbf{C}(t)) \circ \exp(i\mathbf{D}(t))). \quad (10)$$

2.3 新型激励函数

针对复数域中时变问题的求解,将实数域中的非线性激励函数应用到 CVZNN 模型中,提高了模型的收敛速度^[13]。其中,将符号双幂(sign-bi-power, SBP)激励函数应用到 CVZNN 模型中,可以实现有限时间收敛,并能计算出收敛时间上界。因此,可以寻找一种合适的非线性激励函数,将它应用到 NFTZNN-I 模型(7)和 NFTZNN-II 模型(10)中,来加快两种模型的收敛速度、提高其计算精度。因此,选用新型激励函数

$$f(x) = a_1 |x|^k \text{sign}(x) + a_2 x, a_1 > 0, 0 < k < 1, a_2 > 0. \quad (11)$$

式(11)中: a_1 、 k 和 a_2 为激励函数的设计参数; $\text{sign}(\cdot)$ 为符号函数。新型激励函数是单调递增的奇函数,证明如下

证明:新型激励函数为奇函数,根据式(11), $f(0) = 0$, $f(-x) = -a_1 |x|^k \text{sign}(x) - a_2 x$ 。因此 $f(-x) = -f(x)$,故 $f(x)$ 为奇函数。再证明其为单调递增的函数,因为 $f(x)$ 为奇函数,故 $f(x)$ 在 $[-\infty, 0]$ 与 $[0, \infty]$ 的单调性一致,令 $x \in [0, \infty]$,则 $\dot{f}(x) = a_1 k x^{k-1} + a_2 > 0$,故 $f(x)$ 单调递增。综上,新型激励函数是单调递增的奇函数。

根据文献[11],将它应用到两种 NFTZNN 模型时可以保持全局收敛。另外,为进行对比试验,列举出 SBP 激励函数

$$f(x) = \frac{1}{2} |x|^r \text{sign}(x) + \frac{1}{2} |x|^{\frac{1}{r}} \text{sign}(x), 0 < r < 1. \quad (12)$$

式(12)中: r 为 SBP 激励函数的设计参数。

3 理论分析

本节给出四个定理,前两个定理说明当两种 NFTZNN 模型应用单调递增的奇函数,并被用来求解时变复数矩阵方程时,能收敛到理论解。后两个定理说明当两种 NFTZNN 模型应用新型激励函数(11),其在求解时变复数矩阵方程时,能在有限时间内收敛到理论解,并可以计算出收敛时间上界。

3.1 稳定性分析

定理 1 给定光滑的时变复数矩阵 $\mathbf{F}(t)$,如果将单调递增的奇函数 $f(\cdot)$ 应用到 NFTZNN-I 模型(7),那么通过 NFTZNN-I 模型(7)求解的未知矩阵 $\mathbf{Z}(t)$ 从任意初始状态 $\mathbf{Z}(0)$ 开始,最终能收敛到时变复数矩阵方程的理论解 $\mathbf{Z}^*(t)$ 。

证明:根据式(6),将 $\dot{\mathbf{E}}(t)$ 中的实部与虚部分开

$$\begin{cases} \dot{\mathbf{U}}(t) = -\eta \Omega(\mathbf{U}(t)); \\ \dot{\mathbf{V}}(t) = -\eta \Omega(\mathbf{V}(t)). \end{cases} \quad (13)$$

取矩阵 $\mathbf{U}(t)$ 和 $\mathbf{V}(t)$ 中的各个元素,可得:

$$\begin{cases} \dot{u}_{pq}(t) = -\eta f(u_{pq}(t)); \\ \dot{v}_{pq}(t) = -\eta f(v_{pq}(t)). \end{cases} \quad (14)$$

式(14)中: $u_{pq}(t)$ 和 $v_{pq}(t)$ 分别为 $\mathbf{U}(t)$ 和 $\mathbf{V}(t)$ 的第 p 行第 q 列元素, $\dot{u}_{pq}(t)$ 与 $\dot{v}_{pq}(t)$ 是它们的一阶导数; $f(\cdot)$ 是处理复数矩阵各个元素的单调递增的奇函数。定义李亚普诺夫函数 $L_u = \frac{1}{2} u_{pq}^2(t)$,对它求导,可得 $\dot{L}_u = -\eta u_{pq}(t) f(u_{pq}(t))$ 。因为 $f(\cdot)$ 单调递增且为奇函数,所以 $\dot{L}_u \leq 0$ 。根据李亚普诺夫定理[14]可以得出, p, q 取任意值时, $u_{pq}(t)$ 都能全局收敛到零,故矩阵 $\mathbf{U}(t)$ 能全局收敛到零。同理,也可证明出 $\mathbf{V}(t)$ 能全局收敛到零。综上所述,误差矩阵最终能全局收敛到零。因此,当 $\mathbf{E}(t)$ 全局收敛到零时,说明用 NFTZNN-I 模型(7)来求解时变复数矩阵方程时,最终能得到方程的理论解 $\mathbf{Z}^*(t)$ 。此外,根据式(2)中 $\mathbf{E}(t)$ 的定义,求解出的 $\mathbf{Z}(t)$ 与初始状态 $\mathbf{Z}(0)$ 无关。

定理 1 成立。

定理 2 给定光滑的时变复数矩阵 $\mathbf{F}(t)$,如果将单调递增的奇函数 $f(\cdot)$ 应用到 NFTZNN-II 模型(10),那么通过 NFTZNN-II 模型(10)求解的未知矩阵 $\mathbf{Z}(t)$ 从任意初始状态 $\mathbf{Z}(0)$ 开始,最终能收敛到时变复数矩阵方程的理论解 $\mathbf{Z}^*(t)$ 。

证明:根据式(9),可得:

$$\dot{\mathbf{E}}(t) = -\eta(\Omega(\mathbf{C}(t)) \circ \exp(i\mathbf{D}(t))). \quad (15)$$

首先,取 $\dot{\mathbf{E}}(t)$ 中的各个元素

$$\dot{e}_{pq}(t) = -\eta(f(\omega_{pq}(t)) \circ \exp(i\theta_{pq}(t))). \quad (16)$$

式(16)中: $e_{pq}(t)$ 为复数误差矩阵 $\mathbf{E}(t)$ 的第 p 行第 q 列元素,且 $e_{pq}(t) = \omega_{pq} \exp(i\theta_{pq}(t))$; $\dot{e}_{pq}(t)$ 为 $e_{pq}(t)$ 的一阶导数, $\omega_{pq}(t)$ 为它的模, $\theta_{pq}(t)$ 为它的辐角。然后,定义李亚普诺夫函数 $L_e = \frac{1}{2} |e_{pq}(t)|^2 = \frac{1}{2} \omega_{pq}^2(t)$,根据复数性质,式子可以写为 $L_e = \frac{1}{2} |e_{pq}(t)|^2 = \frac{1}{2} e_{pq}(t) \overline{e_{pq}(t)}$,其中 $\overline{e_{pq}(t)}$ 为 $e_{pq}(t)$ 的共轭复数,且 $\overline{e_{ij}(t)} = \omega_{pq}(t) \exp(-i\theta_{pq}(t))$ 。最后,对 L_e 进行求导:

$$\dot{L}_e = \frac{1}{2} (\dot{e}_{pq}(t) \overline{e_{pq}(t)} + e_{pq}(t) \dot{\overline{e_{pq}(t)}}) = -\frac{1}{2} \eta f(\omega_{pq}(t) \exp(i\theta_{pq}(t))) \cdot \omega_{pq}(t) \exp(-i\theta_{pq}(t)) -$$

$$\frac{1}{2}\eta\omega_{pq}(t)\exp(i\theta_{pq}(t)) \cdot f(\omega_{pq}(t))\exp(-i\theta_{pq}(t)) = -\eta\omega_{pq}(t)f(\omega_{pq}(t)) = -\eta|e_{pq}(t)|f(|e_{pq}(t)|). \quad (17)$$

式(17)中: $\overline{\dot{e}_{pq}(t)} = f(\omega_{pq}(t))\exp(-i\theta_{pq}(t))$ 。因为 $f(\cdot)$ 单调递增且为奇函数,并且 $|e_{pq}(t)|$ 恒大于等于零,所以 $\dot{L}_e \leq 0$ 。根据李亚普诺夫定理,可以得出, p, q 取任意值时, $e_{pq}(t)$ 都能全局收敛到零,所以矩阵 $\mathbf{E}(t)$ 最终能全局收敛到零。因此,当 $\mathbf{E}(t)$ 全局收敛到零时,说明用 NFTZNN-II 模型(10)求解时变复数矩阵方程时,最终能得到方程的理论解 $\mathbf{Z}^*(t)$ 。此外,根据式(2)中 $\mathbf{E}(t)$ 的定义,求解出的 $\mathbf{Z}(t)$ 与初始状态 $\mathbf{Z}(0)$ 无关。

定理2成立。

3.2 收敛性分析

定理3 给定光滑的时变复数矩阵 $\mathbf{F}(t)$, 如果将新型激励函数(11)应用到 NFTZNN-I 模型(7), 那么通过 NFTZNN-I 模型(7)求解出的未知矩阵 $\mathbf{Z}(t)$ 从任意初始状态 $\mathbf{Z}(0)$ 开始, 最终能在有限时间后收敛到时变复数矩阵方程的理论解 $\mathbf{Z}^*(t)$ 。NFTZNN-I 模型(7)的收敛时间上界

$$t_1 = \frac{\ln\left[\frac{a_2}{a_1}m(0)^{1-k} + 1\right]}{\eta a_2(1-k)}. \quad (18)$$

式(18)中: $m(0) = \max\{|u_{pq}(0)|, |v_{pq}(0)|\}$; $u_{pq}(0)$ 和 $v_{pq}(0)$ 分别为误差矩阵 $\mathbf{E}(0)$ 第 p 行第 q 列元素的实部与虚部。

证明:根据式(14),可得:

$$\begin{cases} \dot{u}_{pq}(t) = -\eta f(u_{pq}(t)); \\ \dot{v}_{pq}(t) = -\eta f(v_{pq}(t)). \end{cases} \quad (19)$$

设函数 $m(t)$, 在 $t=0$ 时选取 $u_{pq}(t)$ 和 $v_{pq}(t)$ 中较大的一个作为 $m(t)$ 的初值, $m(0) = \max\{|u_{pq}(t)|, |v_{pq}(t)|\}$ 。在任意时刻 t , 取 $u_{pq}(t)$ 和 $v_{pq}(t)$ 中绝对值较大的作为 $m(t)$, 由此可以得到: $-|m(t)| \leq u_{pq}(t) \leq |m(t)|$, $-|m(t)| \leq v_{pq}(t) \leq |m(t)|$ 。换言之, 当 $m(t)$ 等于零时, $\mathbf{U}(t)$ 和 $\mathbf{V}(t)$ 中所有元素都会等于零。因为 $f(\cdot)$ 为单调递增的奇函数, 矩阵 $\dot{\mathbf{U}}(t)$ 和 $\dot{\mathbf{V}}(t)$ 中所有的元素都满足式(19), 故可得 $m(t)$ 的动态变化式为

$$\dot{m}(t) = -\eta f(m(t)). \quad (20)$$

首先定义李亚普诺夫函数 $L(t) = |m(t)|^2$, 对它求导:

$$\begin{aligned} \dot{L}(t) &= -2\eta m(t)f(m(t)) = -2\eta m(t)(a_1 m(t)^k \text{sign}(m(t)) + a_2 m(t)) = \\ &= -2\eta(a_1 m(t)^{k+1} + a_2 m(t)^2) = -2\eta(a_1 L(t)^{\frac{k+1}{2}} + a_2 L(t)). \end{aligned} \quad (21)$$

将式(21)两边同乘 $e^{2\eta a_2 t}$, 化简得:

$$\frac{d(L(t)e^{2\eta a_2 t})}{(L(t)e^{2\eta a_2 t})^{\frac{k+1}{2}}} = -2\eta a_1 e^{(1-k)\eta a_2 t}. \quad (22)$$

最后将微分方程(22)两边同时积分, 解得:

$$L(t) = e^{-2\eta a_2 t} \left(-\frac{a_1}{a_2} e^{(1-k)\eta a_2 t} + \frac{a_1}{a_2} + L(0)^{\frac{1-k}{2}} \right)^{\frac{2}{1-k}}. \quad (23)$$

由于 $L(t) = |m(t)|^2$, 故 $L(t) \geq 0$, 因此式(22)中 $-2\eta a_1 e^{(1-k)\eta a_2 t} \geq 0$, 可以解得:

$$t_1 \leq \frac{\ln\left[\frac{a_2}{a_1}m(0)^{1-k} + 1\right]}{\eta a_2(1-k)}.$$

定理3成立。

定理4 给定光滑的时变复数矩阵 $\mathbf{F}(t)$, 如果将新型激励函数(11)应用到 NFTZNN-II 模型(10), 那么通过 NFTZNN-II 模型(10)求解出的未知矩阵 $\mathbf{Z}(t)$ 从任意初始状态 $\mathbf{Z}(0)$ 开始, 最终能在有限时间后收敛到时变复数矩阵方程的理论解 $\mathbf{Z}^*(t)$ 。NFTZNN-II 模型(10)的收敛时间上界

$$t_2 = \frac{\ln\left[\frac{a_2}{a_1}n(0)^{1-k} + 1\right]}{\eta a_2(1-k)}. \quad (24)$$

式(24)中: $n(0) = \max\{|e_{pq}(0)|\}$, $e_{pq}(0)$ 分别为误差矩阵 $\mathbf{E}(0)$ 的第 p 行第 q 列元素的初始状态。

证明: 定义李亚普诺夫函数 $L(t) = |n(t)|^2$, $n(0)$ 是 $\mathbf{E}(t)$ 所有元素中模数的最大值, 且初始状态 $n(0) = \max\{|e_{pq}(0)|\}$ 。可以得出结论, 当 $n(t)$ 趋于零时, $\mathbf{E}(t)$ 中所有元素都会等于零。因此, 若计算出 $n(t)$ 的收敛时间, 就能估算出 NFTZNN-II 模型(10)的收敛时间上界。根据式(16), 可以得到 $n(t)$ 的动态表达式:

$$\dot{n}(t) = -\eta(f(\omega_{pq}(t)) \circ \exp(i\theta_{pq}(t))). \quad (25)$$

与定理 3 的证明类似, NFTZNN-II 模型(10)的收敛时间上界计算结果为 $t_2 \leq \frac{\ln\left[\frac{a_2}{a_1}n(0)^{1-k} + 1\right]}{\eta a_2(1-k)}$ 。

定理 4 成立。

4 仿真实例

以下通过求解一个时变复数矩阵方程, 并在方程中使用不同的矩阵系数, 得到两个不同的算例, 从而证明使用新型激励函数的 NFTZNN-I 模型(7)和 NFTZNN-II 模型(10)求解时变复数矩阵方程时能保证全局收敛。此外, 为说明误差收敛速度的改进, 将新型激励函数(11)与 SBP 激励函数(12)进行比较, 并计算出了各自的收敛时间上界。

仿真 1 考虑时变复数矩阵方程 $\mathbf{F}(t)\mathbf{Z}(t) = \mathbf{I}$, 已知 $\mathbf{F}(t)$ 和单位矩阵 \mathbf{I} , 求未知矩阵 $\mathbf{Z}(t)$, 矩阵系数如下:

$$\mathbf{F}(t) = \begin{bmatrix} \exp(10it) & -i\exp(-10it) \\ -i\exp(10it) & \exp(-10it) \end{bmatrix}, \mathbf{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

在这种情况下, 计算出 $\mathbf{Z}(t)$ 的理论解为

$$\mathbf{Z}^*(t) = \begin{bmatrix} 0.5\exp(-10it) & 0.5i\exp(-10it) \\ 0.5i\exp(10it) & 0.5\exp(10it) \end{bmatrix}.$$

图 1 为仿真 1 中 NFTZNN-I 模型和 NFTZNN-II 模型中 $\mathbf{Z}(t)$ 的轨迹变化, 描述了使用新型激励函数(11)的 NFTZNN-I 模型(7)和 NFTZNN-II 模型(10), 用来求解时变复数矩阵方程时未知矩阵 $\mathbf{Z}(t)$ 的阶段性变化(两种 NFTZNN 模型中设定 $\eta = 40, a_1 = 0.6, k = 0.5, a_2 = 1$)。图中红色实线表示实际解 $\mathbf{Z}(t)$ 的状态变化, 蓝色虚线表示理论解 $\mathbf{Z}^*(t)$ 的状态变化。图中实线与虚线会在一段时间后重合, 这说明 $\mathbf{Z}(t)$ 从初始状态 $\mathbf{Z}(0)$ 开始(仿真中初步设定初值 $\mathbf{Z}(0) = [1+i, 1; 1+i, 1]$), 最终能在一段时间后收敛到时变复数矩阵方程的理论解 $\mathbf{Z}^*(t)$ 。

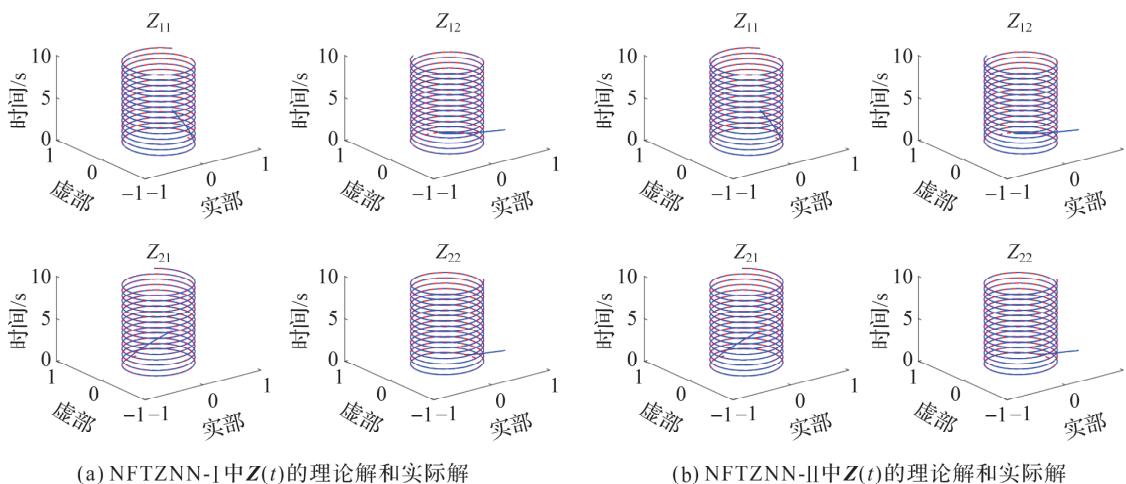


图 1 仿真 1 中 NFTZNN-I 模型和 NFTZNN-II 模型中 $\mathbf{Z}(t)$ 的轨迹变化

Fig. 1 Trajectory change of $\mathbf{Z}(t)$ in NFTZNN-I model and NFTZNN-II model in simulation 1

通常情况下, 实际解与理论解之间存在估计误差, 用 $\|\mathbf{F}(t)\mathbf{Z}(t) - \mathbf{I}\|_F$ 来近似估计实际解与理论解之间的估计误差。由于 SBP 激励函数(11)在求解时变矩阵方程时, 误差的收敛速度较快, 并且能计算出收

敛时间上界,因此选用它来与新型激励函数(11)进行比较。为计算出收敛时间上界,使用特定的初始状态 $\mathbf{Z}(0) = [1+i, 1; 1+i, 1]$, 因此 $\mathbf{E}(0) = \mathbf{F}(0)\mathbf{Z}(0) - \mathbf{I} = [1, 1-i; 2, -i]$, 故式(18)中定义的最大项 $m(0) = 2$ 。将新型激励函数(11)(设定 $a_1 = 0.6, k = 0.5, a_2 = 1$)应用到 NFTZNN-I 模型(7)中,根据式(18),计算出使用新型激励函数的 NFTZNN-I 模型(7)的收敛时间上界 $t_1 = \frac{\ln\left[\frac{a_2}{a_1}m(0)^{1-k} + 1\right]}{\eta a_2(1-k)} \approx 0.061(\text{s})$ 。再将 SBP 激励函数(12)(设定 $\eta = 40, r = 0.3$)应用到 NFTZNN-I 模型(7)中,根据文献[13],计算出其收敛时间上界 $t_2 = \frac{2\|m(0)\|^{1-r}}{\eta(1-r)} \approx 0.116(\text{s})$ (设定 $\eta = 40, r = 0.3$)。图2为仿真1中新型激励函数与 SBP 激励函数的收敛性比较,在 NFTZNN-I 模型(7)中,使用新型激励函数(11)的收敛时间上界是 0.061 s, 小于 SBP 激励函数(12)的收敛时间上界(0.116 s)。使用新型激励函数(11)的实际运行时间是 0.056 s, 小于 SBP 激励函数(12)的实际运行时间(0.073 s), 符合预期。这说明了在 NFTZNN-I 模型(7)中,用新型激励函数(11)来求解时变矩阵方程时,收敛速度更快。

图3为仿真1中设计参数 k 的收敛性比较,描述了用 NFTZNN-I 模型(7)来求解时变复数矩阵方程时,设计参数 k 对误差收敛速度的影响(固定 $\eta = 40, a_1 = 0.6, a_2 = 1$)。从图3中可以看出, k 越接近于零,误差的收敛速度越快。

图4为仿真1中设计参数 η 的收敛性比较,描述了用 NFTZNN-I 模型(7)来求解时变复数矩阵方程时,设计参数 η 对误差收敛速度的影响(固定 $a_1 = 0.6, k = 0.5, a_2 = 1$)。从图4中可以看出, η 越大,误差的收敛速度越快。

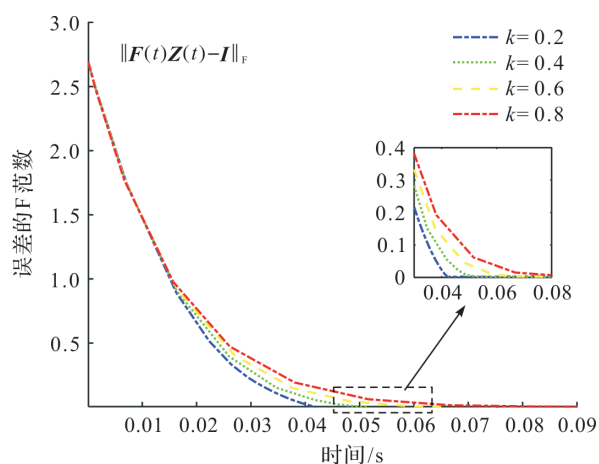


图3 仿真1中设计参数 k 的收敛性比较

Fig. 3 Convergence comparison of design parameter k in simulation 1

仿真2 考虑时变复数矩阵方程 $\mathbf{F}(t)\mathbf{Z}(t) = \mathbf{I}$, 已知 $\mathbf{F}(t)$ 和单位矩阵 \mathbf{I} , 求未知矩阵 $\mathbf{Z}(t)$, 矩阵系数如下:

$$\mathbf{F}(t) = \begin{bmatrix} i\sin(5t) & i\cos(5t) \\ i\cos(5t) & -i\sin(5t) \end{bmatrix}, \mathbf{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

在这种情况下,计算出 $\mathbf{Z}(t)$ 的理论解为

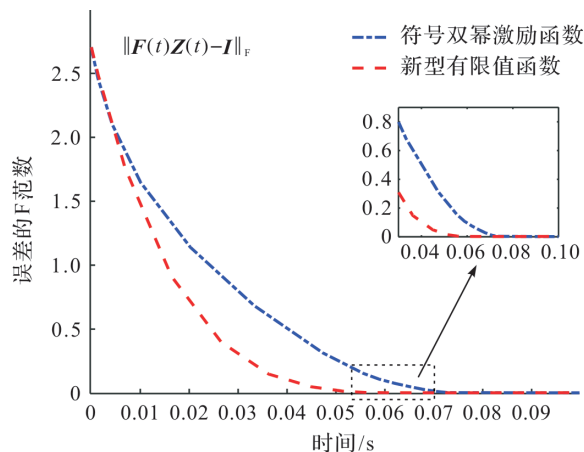


图2 仿真1中新型激励函数与 SBP 激励函数的收敛性比较

Fig. 2 Convergence comparison between new activation function and SBP activation function in simulation 1

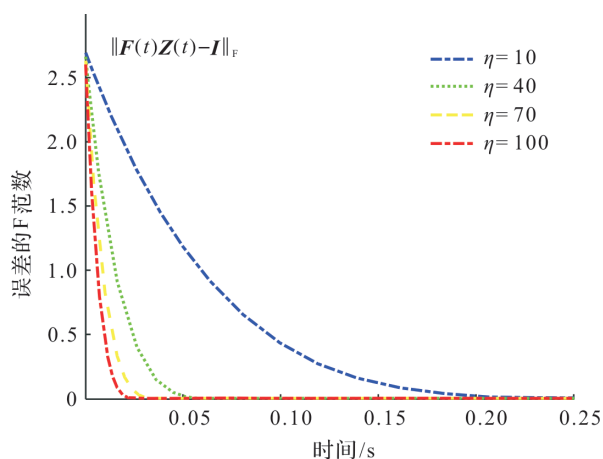


图4 仿真1中设计参数 η 的收敛性比较

Fig. 4 Convergence comparison of design parameter η in simulation 1

$$\mathbf{Z}^*(t) = \begin{bmatrix} -i\sin(5t) & -i\cos(5t) \\ -i\cos(5t) & i\sin(5t) \end{bmatrix}.$$

图 5 为仿真 2 中 NFTZNN-I 模型和 NFTZNN-II 模型中 $\mathbf{Z}(t)$ 的轨迹变化,描述了使用新型激励函数(11)的 NFTZNN-I 模型(7)和 NFTZNN-II 模型(10),用来求解时变复数矩阵方程时未知矩阵 $\mathbf{Z}(t)$ 的阶段性变化(两种 NFTZNN 模型中设定 $\eta = 40, a_1 = 0.6, k = 0.5, a_2 = 1$)。图中实线与虚线会在一段时间后重合,这说明 $\mathbf{Z}(t)$ 从初始状态 $\mathbf{Z}(0)$ 开始(仿真中初步设定初值 $\mathbf{Z}(0) = [3 + 2i, 4 + i; 2 - i, 3 + i]$),最终能在一段时间后收敛到时变复数矩阵方程的理论解 $\mathbf{Z}^*(t)$ 。

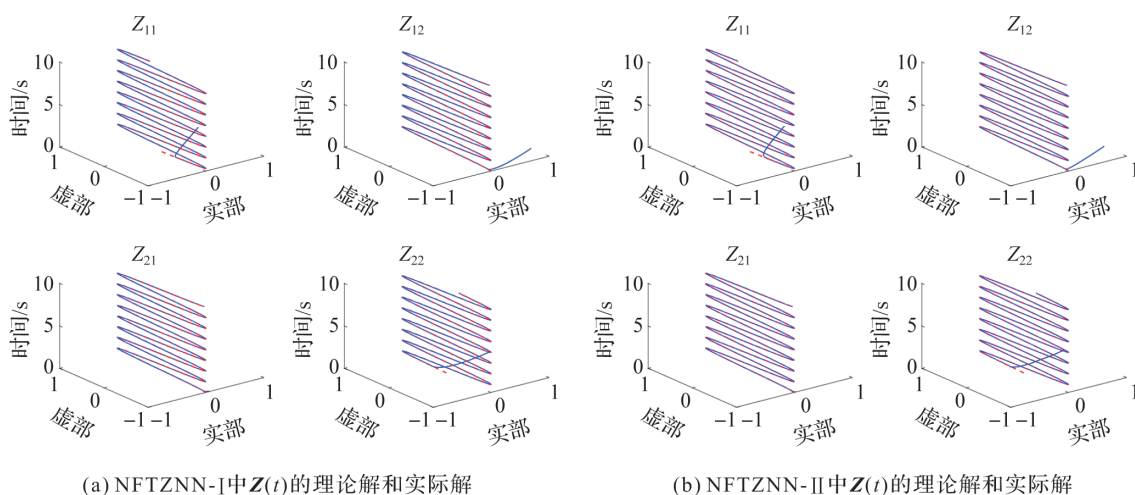


图 5 仿真 2 中 NFTZNN-I 模型和 NFTZNN-II 模型中 $\mathbf{Z}(t)$ 的轨迹变化

Fig. 5 Trajectory change of $\mathbf{Z}(t)$ in NFTZNN-I model and NFTZNN-II model in simulation 2

与仿真实验 1 类似,将新型激励函数(11)与符号双幂激励函数(12)进行误差的收敛性比较。为计算出收敛时间上界,使用特定的初始状态 $\mathbf{Z}(0) = [3 + 2i, 4 + i; 2 - i, 3 + i]$,因此 $\mathbf{E}(0) = \mathbf{F}(0)\mathbf{Z}(0) - \mathbf{I} = [2i, -1 + 3i; -2 + 3i, -2 + 4i]$,故式(24)中定义的最大项 $m(0) = 2\sqrt{5}$ 。将新型激励函数(11)(设定 $a_1 = 0.6, k = 0.5, a_2 = 1$)应用到 NFTZNN-II 模型(10)中,根据式(24)计算出使用新型激励函数的 NFTZNN-II 模型(10)的收敛时间上界 $t_1 = \frac{\ln\left[\frac{a_2}{a_1}m(0)^{1-k} + 1\right]}{\eta a_2(1-k)} \approx 0.076(\text{s})$ 。再将 SBP 激励函数(12)(设定 $\eta = 40, r = 0.3$)应用到 NFTZNN-II 模型(10)中,根据文献[13],计算出其收敛时间上界 $t_2 = \frac{2\|m(0)\|^{1-r}}{\eta(1-r)} \approx 0.204(\text{s})$ (设定 $\eta = 40, r = 0.3$)。

图 6 为仿真 2 中新型激励函数与 SBP 激励函数的收敛性比较,在 NFTZNN-II 模型(10)中,使用新型激励函数(11)的收敛时间上界是 0.076 s,小于 SBP 激励函数(12)的收敛时间上界(0.204 s)。使用新型激励函数(11)的实际运行时间是 0.072 s,小于 SBP 激励函数(12)的实际运行时间(0.079 s),符合预期。这说明在 NFTZNN-II 模型(10)中,用新型激励函数(11)来求解时变矩阵方程时,收敛速度更快。

图 7 为仿真 2 中设计参数 k 的收敛性比较,描述了用 NFTZNN-II 模型(10)求解时变复数矩阵方程时,设计参数 k 对误差收敛速度的影响(固定 $\eta = 40, a_1 = 0.6, a_2 = 1$)。从图 7 中可以看出, k 越接近于零,误差的收敛速度越快。

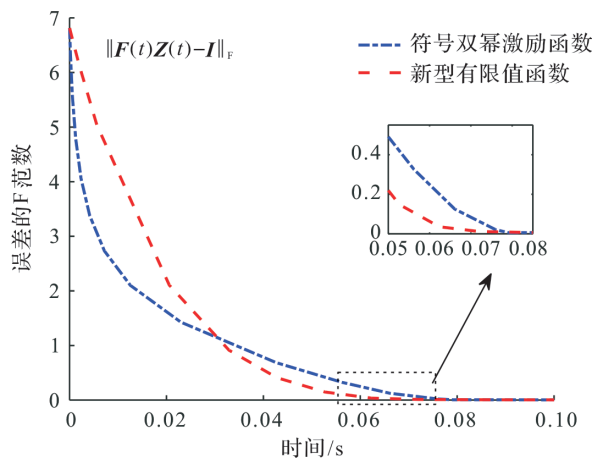


图 6 仿真 2 中新型激励函数与 SBP 激励函数的收敛性比较

Fig. 6 Convergence comparison between new activation function and SBP activation function in simulation 2

图8为仿真2中设计参数 η 的收敛性比较,描述了用NFTZNN-II模型(10)求解时变复数矩阵方程时,设计参数 η 对误差收敛速度的影响(固定 $a_1 = 0.6, k = 0.5, a_2 = 1$)。从图8中可以看出, η 越大,误差的收敛速度越快。

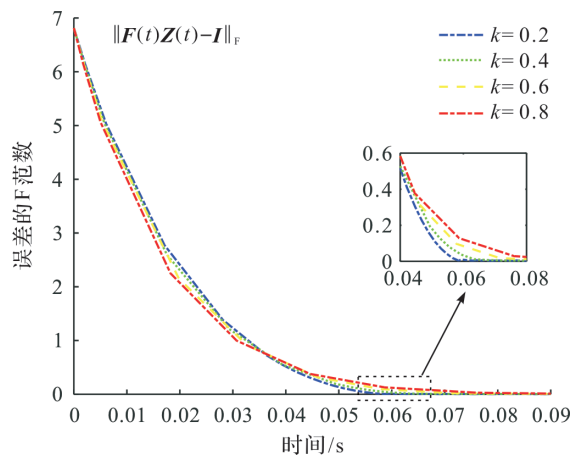


图7 仿真2中设计参数 k 的收敛性比较

Fig. 7 Convergence comparison of design parameter k in simulation 2

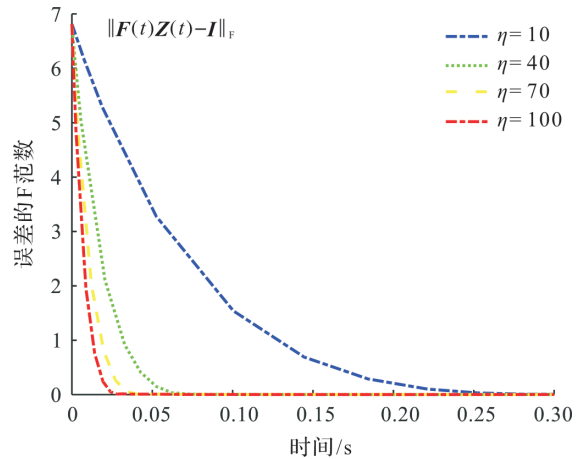


图8 仿真2中设计参数 η 的收敛性比较

Fig. 8 Convergence comparison of design parameter η in simulation 2

5 结 语

本研究针对时变复数矩阵方程的求解,提出了两种新型有限时间归零神经网络(NFTZNN)模型。将一种新型激励函数(11)应用到NFTZNN-I模型(7)和NFTZNN-II模型(10)中,加快两种模型的收敛速度,提高其计算精度。与传统的SBP激励函数(11)相比,应用新型激励函数的两种NFTZNN模型在求解时变复数矩阵方程时,收敛速度更快,收敛时间上界更小。理论分析和数值仿真证明了方法的有效性和可行性。但是,因为新型激励函数是非线性函数,所以使用新型激励函数的NFTZNN模型求解时变复数矩阵方程时,与传统的CVZNN模型相比具有更高的复杂度。未来的研究中可能会考虑噪声对两个NFTZNN模型的影响。

参考文献:

- [1] ZHANG J A, HUANG X, SUZUKI H, et al. Gaussian approximation based interpolation for channel matrix inversion in MIMO-OFDM systems[J]. IEEE Transactions on Wireless Communications, 2013, 12(3): 1407.
- [2] DAHIPHALE V, BANSOD G, ZAMBARE A, et al. Design and implementation of various datapath architectures for the ANU lightweight cipher on an FPGA[J]. Frontiers of Information Technology & Electronic Engineering, 2020, 21(4): 14.
- [3] 于树友, 孟凌宇, 许芳, 等. 基于三步法的机械臂轨迹跟踪控制[J]. 控制理论与应用, 2020, 37(5): 1034.
- [4] TANG C, TAO Y, CHEN Y, et al. Approximate iteration detection and precoding in massive MIMO[J]. China Communications, 2018, 15(5): 183.
- [5] FORSYTHE G E, LEIBLER R A. Matrix inversion by a Monte Carlo method[J]. Mathematics of Computation, 1950, 4(31): 127.
- [6] 崔伟成, 杨利斌, 桑德一. 一种局部特征尺度分解局部均值点求取方法[J]. 组合机床与自动化加工技术, 2020(9): 47.
- [7] CAO Y, SAMIDURAI R, SRIRAMAN R. Robust passivity analysis for uncertain neural networks with leakage delay and additive time-varying delays by using general activation function[J]. Mathematics and Computers in Simulation, 2019, 155: 57.
- [8] XU C. Local and global Hopf bifurcation analysis on simplified bidirectional associative memory neural networks with multiple delays[J]. Mathematics and Computers in Simulation, 2018, 149: 69.

- [9] 周日辉. 基于计算机软件技术的模式识别系统智能开发[J]. 电子技术与软件工程, 2019(21):37.
- [10] DAI J, JIA L, XIAO L. Design and analysis of two prescribed-time and robust ZNN models with application to time-variant stein matrix equation[J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 32(4): 1668.
- [11] 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.
- [12] PANDE A, Goel V. Complex-valued neural network in image recognition: a study on the effectiveness of radial basis function[J]. World Academy of Science, Engineering and Technology, 2007, 26:220.
- [13] LI S, LI Y. Nonlinearly activated neural network for solving time-varying complex Sylvester equation[J]. IEEE transactions on cybernetics, 2013, 44(8):1400.
- [14] XIAO L, ZHANG Y, LI K, et al. A novel recurrent neural network and its finite-time solution to time-varying complex matrix inversion[J]. Neurocomputing, 2019, 331:487.
- [15] 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):767.

~~~~~

(上接第 390 页)

- [12] 郭智勇, 吴攀, 徐伟, 等. 一种三轴双向补偿式磁场测量装置: CN201920265718. 1[P]. 2020-03-17.
- [13] 金贵. 磁感应相位移技术检测脑出血的实验研究[D]. 重庆: 第三军医大学, 2014.
- [14] 何哲玺, 王巍, 王学锋, 等. 一种用于微型核磁共振陀螺仪的磁补偿线圈结构件: CN201510206416. 3[P]. 2015-08-05.
- [15] HÄMÄLÄINEN M, HARI R, ILMONIEMI R J, et al. Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain[J]. Reviews of Modern Physics, 1993, 65(2):413.
- [16] KIRSCHVINK J L. Uniform magnetic fields and double-wrapped coil systems: improved techniques for the design of bioelectromagnetic experiments[J]. Bioelectromagnetics, 1992, 13(5):401.
- [17] NOURI N, PLASTER B. Comparison of magnetic field uniformities for discretize and finite-sized standard  $\cos\theta$ , solenoidal, and spherical coils[J]. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 2013, 723:30.
- [18] WANG J, SONG X, LE Y, et al. Design of self-shielded uniform magnetic field coil via modified pigeon-inspired optimization in miniature atomic sensors[J]. IEEE Sensors Journal, 2021, 21(1):315.
- [19] PIERRO F, ZHAO Z, OWEN C M, et al. Finite-element analysis of the strain distribution due to bending in a REBCO coated conductor for canted cosine theta dipole magnet applications[J]. IEEE Transactions on Applied Superconductivity, 2019, 29(5):260.
- [20] 程军. 载流圆线圈磁场的 MATLAB 数值计算[J]. 物理通报, 2018(5):19.