Neural Networks 98 (2018) 102–113

Contents lists available at ScienceDirect

Neural Networks

journal homepage: www.elsevier.com/locate/neunet

Nonlinear recurrent neural networks for finite-time solution of general time-varying linear matrix equations ☆

Lin Xiao [a,*], Bolin Liao [a], Shuai Li [b], Ke Chen [c]

[a] College of Information Science and Engineering, Jishou University, Jishou 416000, China
[b] Department of Computing, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong
[c] Department of Signal Processing, Tampere University of Technology, Finland

ARTICLE INFO

Article history:
Received 11 July 2017
Received in revised form 25 September 2017
Accepted 16 November 2017
Available online 2 December 2017

Keywords:
Nonlinear recurrent neural networks
General time-varying linear matrix equations
Finite-time convergence
Nonlinear activation functions

ABSTRACT

In order to solve general time-varying linear matrix equations (LMEs) more efficiently, this paper proposes two nonlinear recurrent neural networks based on two nonlinear activation functions. According to Lyapunov theory, such two nonlinear recurrent neural networks are proved to be convergent within finite-time. Besides, by solving differential equation, the upper bounds of the finite convergence time are determined analytically. Compared with existing recurrent neural networks, the proposed two nonlinear recurrent neural networks have a better convergence property (i.e., the upper bound is lower), and thus the accurate solutions of general time-varying LMEs can be obtained with less time. At last, various different situations have been considered by setting different coefficient matrices of general time-varying LMEs and a great variety of computer simulations (including the application to robot manipulators) have been conducted to validate the better finite-time convergence of the proposed two nonlinear recurrent neural networks.

1. Introduction

Linear matrix equations (LMEs) are omnipresent in control theory (De Queiroz, Dawson, Nagarkatti, & Zhang, 2012; Zhou, Duan, & Lin, 2011), optimization (Jin & Zhang, 2015a), signal processing (Xia, Sun, & Zheng, 2012), robotics (Zhang, Li, Zhang, Luo, & Li, 2015), and multi-agent systems (Zhang, Li, Qu, & Lewis, 2015). In mathematics, the general LME is expressed in the form:

$$\sum_{k=1}^{p} A_k X B_k = C \in \mathbb{R}^{n\times m}, \qquad (1)$$

where $A_k \in \mathbb{R}^{n\times n}$ and $B_k \in \mathbb{R}^{m\times m}$ (with $k = 1, 2, \ldots, p$) represent coefficient matrices; $C \in \mathbb{R}^{n\times m}$ represents an arbitrary coefficient; and $X(t) \in \mathbb{R}^{n\times m}$ represents an unknown matrix that needs to be solved. The well-known matrix inversion equation, Lyapunov equation, and Sylvester equation can be deemed as some particular cases of general LMEs when coefficient matrices are set as different values. For example, if $p = 2$, and $B_1 = A_2 = I$ (with I being identity matrix of conforming dimensions), then Eq. (1) is called the Sylvester equation. The Lyapunov equation can also be derived, as long as $p = 2$, $B_1 = A_2 = I$, and $B_2 = A_1^{\mathrm{T}}$. It is worth pointing out that these LMEs have an important role in the stability analysis of linear control systems and also participate in the theoretical developments of some nonlinear systems. Therefore, there are tremendous amount of methods (Benner & Breiten, 2014; Ding & Chen, 2005; Ding & Zhang, 2014; Hajarian, 2016; Peng, Hu, & Zhang, 2005; Simoncini, 2016; Wu & Chang, 2016) that have been developed and studied for the real-time solution of LMEs, including matrix inversion equation, Sylvester equation, Lyapunov equation, and so on.

Most of methods for LMEs appeared in the literature can be grossly classified into two categories: the first is the serial method (e.g., iterative algorithms), and the second one is the parallel method (e.g., neural networks). For the serial method, the iterative algorithm is one of the most typical representative. For example, an iteration method is constructed in Peng et al. (2005) to solve the linear matrix equation $AXB = C$. By this iteration method, the solvability of the equation $AXB = C$ can be determined automatically. In Wu and Chang (2016), two novel iterative algorithms are presented to solve the Lyapunov matrix equations appearing in discrete-time periodic linear systems. In Benner and Breiten (2014), some optimality results for the approximation of large-scale matrix equations are discussed. In particular, this includes the special case of Lyapunov and Sylvester equations.

☆ This work was supported by the National Natural Science Foundation of China under grant 61503152, the Natural Science Foundation of Hunan Province, China under grants 2016JJ2101 and 2017JJ3258, and the National Natural Science Foundation of China under grants 61563017, 61363073, 61662025, and 61561022.
* Corresponding author.
E-mail addresses: xiaolin860728@163.com, xiaolin@jsu.edu.cn (L. Xiao).

https://doi.org/10.1016/j.neunet.2017.11.011
0893-6080/

In Ding and Chen (2005), Ding and Zhang (2014) and Hajarian (2016), some gradient-based iterative algorithms are proposed for solving LMEs, including Sylvester equation and the coupled matrix equations. More computational methods for linear matrix equations are reviewed in Simoncini (2016). It is worth pointing out that iterative algorithms for general LMEs usually lead to the high computational complexity due to their inherent drawbacks. Furthermore, when iterative algorithms are applied to general time-varying LMEs, they should be finished within every sampling period. If not, the algorithms will not accomplish the calculation in a sampling period, which indicates the failure of iterative algorithms.

For the parallel method, the recurrent neural network (RNN) is one of the most typical representative, which found a great of applications in the past decades (Jin & Zhang, 2015b; Liao, Zhang, & Jin, 2016; Liu, Li, Tong, & Chen, 2016; Yu, Shi, Dong, Chen, & Lin, 2015). Different from many conventional iterative algorithms, RNNs can be realizable on specific parallel and distributed hardware architectures (Hosseini, 2016; Peng, Wu, Song, & Shi, 2017; Qin, Liu, Xue, & Wang, 2016; Song, Yan, Zhao, & Liu, 2016; Tu, Cao, & Hayat, 2016; Xiao, 2017a, b). This can highly enlarge utility of current RNNs towards various potential application domains and toward high-performance computing. As competitive computational tools, RNNs play an important role in solving general LMEs (Li, Chen, & Liu, 2013; Xiao, 2015; Xiao & Liao, 2016; Xiao & Lu, 2015; Xiao & Zhang, 2011, 2012; Yi, Chen, & Lan, 2013; Yi, Chen, & Lu, 2011; Zhang, Chen, Li, Yi, & Zhu, 2008; Zhang & Ge, 2005; Zhang, Jiang, & Wang, 2002). For example, gradient-based RNNs are developed for solving time-invariant LMEs, and the solution errors can decrease to zero when time goes to infinity (Yi et al., 2013, 2011; Zhang et al., 2008). In contrast, for the time-varying case, the solution errors are always oscillating since the velocity compensation of time-varying coefficients is not considered. Zhang neural network (ZNN) is thus proposed to solve various time-varying problems, which has a significant improvement in convergence property (Xiao & Zhang, 2011, 2012; Zhang & Ge, 2005; Zhang et al., 2002). In comparison with the gradient-based RNNs, ZNN can avoid the oscillation of the solution errors effectively and can converge to the theoretical solutions of time-varying LMEs exponentially. After that, a novel nonlinear function (termed the sign-bi-power function) is presented to activate ZNN for time-varying LMEs such that its convergence property can achieve the finite time convergence (Li, Chen et al., 2013; Xiao & Liao, 2016). However, owing to the deep investigation, the sign-bi-power activation function has a relatively redundant formulation for finite-time convergence. Therefore, the convergence speed of ZNN can be further improved by optimizing the structure of the sign-bi-power activation function.

Different from the previous study, in this paper, a general framework of nonlinear RNNs for general time-varying LMEs is first proposed on basis of ZNN with the global stability ensured. Then, in order to expedite the finite-time convergence, two novel nonlinear functions are presented to activate the general nonlinear RNN. As compared to the sign-bi-power activation function, the proposed two novel nonlinear functions are of simpler structure, and are convenient for computer simulations as well as hardware implementations. More importantly, the resultant two nonlinear RNNs possess a better convergence property, when compared to the existing neural networks. The main contributions of this work can be summed up as below.

(1) This paper focuses on general time-varying linear matrix equations (LMEs) solving, instead of the specific matrix equation (e.g., Sylvester equation) or general time-invariant LMEs solving.
(2) A general framework of nonlinear recurrent neural networks (RNNs) is proposed for general time-varying LMEs with the global stability ensured.
(3) Two novel nonlinear functions are presented to activate the general nonlinear RNN model, and thus two specific nonlinear RNNs are proposed for general time-varying LMEs solving. In addition, their theoretical analyses are given out to ensure the better finite-time convergence property.
(4) The existing neural networks are applied to general time-varying LMEs solving for comparative purposes, and the extensive simulation results (including the application to robot manipulators) validate the effectiveness and better finite-time convergence of the proposed two nonlinear RNNs for general time-varying LMEs.

2. Problem formulation and model

In this part, we consider the following general time-varying linear matrix equation (LME), which incorporates time-invariant LME (1) as a special case (Li, Chen et al., 2013; Yi et al., 2011; Zhang & Ge, 2005; Zhang et al., 2002):

$$\sum_{k=1}^{p} A_k(t)X(t)B_k(t) = C(t) \in \mathbb{R}^{n\times m}, \tag{2}$$

where $A_k(t) \in \mathbb{R}^{n\times n}$, $B_k(t) \in \mathbb{R}^{m\times m}$ (with $k = 1, 2, \ldots, p$), and $C(t) \in \mathbb{R}^{n\times m}$ represent time-varying coefficient matrices. In addition, $A_k(t) \in \mathbb{R}^{n\times n}$ and $B_k(t) \in \mathbb{R}^{m\times m}$ satisfies the condition of the unique solution of (2) (Simoncini, 2016). The aim of this work is to find the unknown time-varying matrix $X(t) \in \mathbb{R}^{n\times m}$ within finite time by designing the nonlinear recurrent neural networks such that the above general time-varying LME (2) holds.

2.1. The model

In the literature, Zhang neural network (ZNN) has been proposed to solve various time-varying problems broadly encountered in scientific and engineering areas (Xiao & Zhang, 2011, 2012; Zhang et al., 2008; Zhang & Ge, 2005; Zhang et al., 2002). Considering the advantages of the exponential convergence, the design method of ZNN is first extended to solve the general time-varying linear matrix equation (2). The realization can be summarized and listed in the following procedure (Xiao & Zhang, 2011, 2012; Zhang et al., 2008; Zhang & Ge, 2005).

First, an indefinite matrix-valued error function $\mathfrak{E}(t)$ is defined based on the formulation of the general time-varying linear matrix equation (2):

$$\mathfrak{E}(t) = \sum_{k=1}^{p} A_k(t)X(t)B_k(t) - C(t) \in \mathbb{R}^{n\times m}, \tag{3}$$

of which each element is indefinite. It can be negative, positive, zero, or even lower-unbounded.

Second, the following evolution formula is adopted for $\mathfrak{E}(t)$ such that $\lim_{t\to\infty}\mathfrak{E}(t) = 0$:

$$\frac{\mathrm{d}\mathfrak{E}(t)}{\mathrm{d}t} = -\gamma\mathcal{F}(\mathfrak{E}(t)), \tag{4}$$

where design parameter $\gamma > 0$ is the scaling factor to adjust the convergence rate of $\lim_{t\to\infty}\mathfrak{E}(t) = 0$, and $\mathcal{F}(\cdot)$ stands for the nonlinear activation function array.

At last, expanding the evolution formula (4) by substituting (3) into (4), we obtain the general framework of the nonlinear