

A complex neural network algorithm for computing the largest sum of real part and imaginary part of eigenvalues and the corresponding eigenvector of a real normal matrix

Hang Tan[†], Li-Ping Wan, Rong Ye, Xue-Song Liang and Zhao-Yao Wu

*College of physics and engineering, Chengdu Normal University,
Chengdu, 611130, China*

[†]E-mail: hangtancdu@163.com

www.cdnu.edu.cn

In this study, we proposed a novel complex neural network algorithm, which extends the neural networks based approaches that can asymptotically compute the largest modulus of eigenvalues and the corresponding eigenvector to the case of directly computing the largest sum of real part and imaginary part of eigenvalues and the corresponding eigenvectors of a real normal matrix. The proposed neural network algorithm is described by a group of complex differential equations. And the algorithm has parallel processing ability in an asynchronous manner and could achieve high computing capability. This paper also provides a rigorous mathematical proof for its convergence for a more clear understanding of network dynamic behaviors relating to the computation of the eigenvector and the eigenvalue. Numerical example showed that the proposed algorithm has good performance for a general real normal matrix.

Keywords: Complex neural network; Real normal matrix; Maximum sum of real part and imaginary part; Eigenvalue; Eigenvector.

1. Introduction

The neural network technology has been used to extract the eigen-pairs of real matrices is proposed about 1980's [1], and then motivated very broad interests among engineering and theoretical researches [2-8]. In 1995, Luo et al. [2,3] presented another neural network algorithm for computing not only modulus largest eigenvalue but also modulus smallest eigenvalue and their corresponding eigenvectors of real symmetric matrices. In recent years, some adaptive generalized eigen-pairs extraction algorithms of Hermite matrices have been developed by some authors in [7, 8].

About some years ago, Liu et al. [4] proposed a simpler neural network algorithm for extracting the eigenvector corresponding to the modulus maximum eigenvalue of a real symmetric matrix, in the paper [5], they extended this model to extract the imaginary part of the eigenvalue from the maximum imaginary part and the real part of the eigenvalue from the maximum real part of a general

real matrix. In the present paper, we also along with this way, however, we intend try to extract the eigenvalue and the corresponding eigenvector from the largest sum of the real part and the imaginary part of eigenvalues of a general real normal matrix.

The rest of this paper is organized as follows: In section 2, we proposed the novel complex domain neural network model for our purpose and the neural network convergence analysis is presented in section 3. The numerical example given in section 4 and in the last section we summarized this paper.

2. The Novel Complex Neural Network Algorithm

We also consider the following real neural networks model in [1]:

$$\frac{dv(t)}{dt} = Av(t) - v^T(t)Av(t)v(t), \quad (1)$$

In which $v(t) \in R^n$ are the n dimension real column vectors that denote the states of neurons. In [5], authors proposed a new method to compute the largest real part and the largest imaginary part of eigenvalues from a real matrix in complex domain. In this paper, we intended to extract the eigenvalue and the corresponding eigenvector from the largest sum of the real part and the imaginary part of eigenvalues of a real normal matrix by modifying the original model.

In Eq. (1), we instead of the matrix A by A' :

$$A' = \begin{pmatrix} A & A \\ -A & A \end{pmatrix}.$$

Where A is a general real matrix, $v(t) \in R^{2n}$, and assume that $v(t)^T = [x(t)^T, y(t)^T]$, and Let $z(t) = x(t) + y(t)i$, so Eq. (1) can be converted into the following forms:

$$\frac{dz(t)}{dt} = (1-i)Az(t) - \left[\frac{z^T A \bar{z} + \bar{z}^T A z}{2} - \frac{z^T A \bar{z} - \bar{z}^T A z}{2i} \right] z(t). \quad (2)$$

Eq. (2) is the complex neural network model in our study, where $z(t) \in C^n$. The algorithm can be used to extract the eigenvector and the corresponding eigenvalue, which has largest real part and imaginary part sum of a real normal matrix.

Assume that all eigenvalues of real normal matrix A are denoted as $\lambda_1^R + \lambda_1^I i, \lambda_2^R + \lambda_2^I i, \dots, \lambda_n^R + \lambda_n^I i$, where $\lambda_k^R, \lambda_k^I \in R, k = 1, 2, \dots, n$, and

the corresponding normalized complex eigenvectors are denoted as S_1, S_2, \dots, S_n , obviously, they construct a normalized orthogonal basis in $C^{n \times n}$.

Assume that

$$z(t) = \sum_{k=1}^n z_k(t) S_k = \sum_{k=1}^n [x_k(t) + iy_k(t)] S_k, \quad (3)$$

Where $z_k(t) = x_k(t) + iy_k(t)$ denote the projection of $z(t)$ in the direction along with S_k .

Theorem 2.1. Let $z_k(t) = x_k(t) + iy_k(t)$ denote the projection of $z(t)$ in the direction along with S_k , the analytic expression of $|z(t)|^2$ can be read as

$$|z(t)|^2 = \frac{\sum_{k=1}^n \exp[2(\lambda_k^R + \lambda_k^I)t] |z_k(0)|^2}{1 + 2 \sum_{j=1}^n (\lambda_j^R + \lambda_j^I) |z_j(0)|^2 \int_0^t \exp[2(\lambda_j^R + \lambda_j^I)\tau] d\tau}. \quad (4)$$

Proof. We consider the projection of $z(t)$ onto S_k . Substituting Eq. (3) into Eq. (2), by use of $AS_k = \lambda_k S_k$, one can get

$$\begin{aligned} \sum_{k=1}^n \left[\frac{dx_k(t)}{dt} + i \frac{dy_k(t)}{dt} \right] S_k &= (1-i) \sum_{k=1}^n \lambda_k (x_k + iy_k) S_k \\ &- \left[\frac{1}{2} \left\{ \sum_{j=1}^n [\lambda_j + \bar{\lambda}_j] |z_j|^2 \right\} + \frac{1}{2i} \left\{ \sum_{j=1}^n [\lambda_j - \bar{\lambda}_j] |z_j|^2 \right\} \right] \sum_{k=1}^n (x_k + iy_k) S_k. \end{aligned} \quad (5)$$

Along with S_k , Substitute of $\lambda_j = \lambda_j^R + i\lambda_j^I$, $\bar{\lambda}_j = \lambda_j^R - i\lambda_j^I$ and $\lambda_k = \lambda_k^R + i\lambda_k^I$ into Eq. (5). After separating the real part and the imaginary part, and consider that $\frac{d|z_k|^2}{dt} = \frac{d(\bar{z}_k z_k)}{dt} = 2x_k \frac{dx_k}{dt} + 2y_k \frac{dy_k}{dt}$ we have

$$\frac{d|z_k(t)|^2}{dt} = 2|z_k|^2 (\lambda_k^R + \lambda_k^I) - 2 \sum_{j=1}^n (\lambda_j^R + \lambda_j^I) |z_j|^2 |z_k|^2. \quad (6)$$

From Eq. (6), if $|z_k(t)|^2 \neq 0$, $|z_r(t)|^2 \neq 0$, we have

$$\frac{|z_k|^2}{|z_r|^2} = \frac{|z_k(0)|}{|z_r(0)|^2} \exp\{2[(\lambda_k^R + \lambda_k^I) - (\lambda_r^R + \lambda_r^I)]t\}. \quad (7)$$

From Eq. (6), we can directly write it as follows:

$$\frac{1}{|z_k|^4} \frac{d}{dt} |z_k|^2 = 2(\lambda_k^R + \lambda_k^I) \frac{1}{|z_k|^2} - 2 \sum_{j=1}^n (\lambda_j^R + \lambda_j^I) \frac{|z_j|^2}{|z_k|^2}. \quad (8)$$

One gets

$$\frac{d}{dt} \frac{\exp[2(\lambda_k^R + \lambda_k^I)t]}{|z_k(t)|^2} = 2 \sum_{j=1}^n (\lambda_j^R + \lambda_j^I) \frac{|z_j(0)|^2}{|z_k(0)|^2} \exp[2(\lambda_j^R + \lambda_j^I)t]. \quad (9)$$

The integral on both sides of Eq. (15) from 0 to t can reads, Direct calculation reads

$$|z_k(t)|^2 = \frac{\exp[2(\lambda_k^R + \lambda_k^I)t] \times |z_k(0)|^2}{1 + 2 \sum_{j=1}^n (\lambda_j^R + \lambda_j^I) |z_j(0)|^2 \int_0^t \exp[2(\lambda_j^R + \lambda_j^I)\tau] d\tau},$$

Therefore one gets Eq. (4) \square

3. Convergence Analysis

If an equilibrium vector of the neural network Eq. (2) exists, let ξ denote it, and there exists

$$\xi = \lim_{t \rightarrow \infty} z(t) \quad (10)$$

Theorem 3.1. *If A is not a zero matrix, then $|\xi| \neq 0$.*

Proof. Obviously, from the theorem 2.1, we have

$$|\xi| = \lim_{t \rightarrow \infty} |z(t)| = \lim_{t \rightarrow \infty} \sqrt{\frac{\sum_{k=1}^n \exp[2(\lambda_k^R + \lambda_k^I)t] \times |z_k(0)|^2}{1 + 2 \sum_{j=1}^n (\lambda_j^R + \lambda_j^I) |z_j(0)|^2 \int_0^t \exp[2(\lambda_j^R + \lambda_j^I)\tau] d\tau}},$$

$|\xi| = 0$ implies that all eigenvalues of A are 0, so A should be a zero matrix. \square

Theorem 3.2. *If each eigenvalue of A satisfies that $\lambda_k^R + \lambda_k^I < 0$, ($k = 1, 2, \dots, n$), then $|\xi| = 0$.*

Proof. From the theorem 2.1 and the theorem 3.1, we have that $\sum_{k=1}^n \exp[2(\lambda_k^R + \lambda_k^I)t] \times |z_k(0)|^2 = 0$, so $|\xi| = 0$. In this case, the model Eq. (2) can't work correctly in this case. \square

Theorem 3.3. *Denote $(\lambda_m^R + \lambda_m^I) = \max_{1 \leq k \leq n} (\lambda_k^R + \lambda_k^I)$. If $(\lambda_m^R + \lambda_m^I) \geq 0$, then $\xi^T \bar{\xi} = 1$.*

Proof. From Eq. (9) and the theorem 2.1, we have

$$\begin{aligned} \xi^T \bar{\xi} &= \lim_{t \rightarrow \infty} |z(t)|^2 = \lim_{t \rightarrow \infty} \frac{\sum_{k=1}^n \exp[2(\lambda_k^R + \lambda_k^I)t] |z_k(0)|^2}{1 + 2 \sum_{j=1}^n (\lambda_j^R + \lambda_j^I) |z_j(0)|^2 \int_0^t \exp[2(\lambda_j^R + \lambda_j^I)\tau] d\tau} \\ &= \lim_{t \rightarrow \infty} \frac{|z_m(0)|^2 + \sum_{k=1, k \neq m}^n \exp[2(\lambda_k^R + \lambda_k^I - \lambda_m^R - \lambda_m^I)t] |z_k(0)|^2}{\exp[-2(\lambda_m^R + \lambda_m^I)t] \{1 + 2 \sum_{j=1, j \neq m}^n (\lambda_j^R + \lambda_j^I) |z_j(0)|^2 \int_0^t \exp[2(\lambda_j^R + \lambda_j^I)\tau] d\tau\} + |z_m(0)|^2 \{1 - \exp[-2(\lambda_m^R + \lambda_m^I)t]\}} \\ &= 1 \end{aligned} \quad \square$$

Theorem 3.4. For Eq. (3), if there are no exist two or more different eigenvalues shared the same largest sum of real part and the imaginary part of eigenvalues then $\lim_{t \rightarrow \infty} \frac{\bar{z}^T(t) A z(t)}{\bar{z}^T(t) z(t)}$ will converge to the eigenvalue that has the largest sum of the real part and the imaginary part, and the corresponding eigenvector is $\xi = \lim_{t \rightarrow \infty} z(t)$.

Proof. From the theorem 2.1, one could easily to get that $\lim_{t \rightarrow \infty} \bar{z}^T(t) A z(t) = \lambda_m \lim_{t \rightarrow \infty} \bar{z}^T(t) z(t)$, $\lim_{t \rightarrow \infty} \bar{z}^T(t) [A z(t) - \lambda_m z(t)] = 0$, from the theorem 3.1, we know that $\lim_{t \rightarrow \infty} \bar{z}^T(t) \neq 0$, so we have $\lim_{t \rightarrow \infty} [A z(t) - \lambda_m z(t)] = 0$, directly we get $A \lim_{t \rightarrow \infty} z(t) = \lambda_m \lim_{t \rightarrow \infty} z(t)$, so $\xi = \lim_{t \rightarrow \infty} z(t)$ is the eigenvector corresponding to the eigenvalue λ_m , which has the largest sum of the real part and the imaginary part. \square

4. Simulation Results

Consider the above 6×6 real artificial normal matrix A , the six eigenvalues of A are $11.0000 \pm 9.0000i, -15.0000, -10 \pm 4.0000i, -3.0000$, respectively. We use the stochastic initial value $z(0) = (1, 1, 1, 1, 1, 1)^T$ for running the complex neural network algorithm Eq. (2) to obtain the largest sum of real part and imaginary part of eigenvalues of matrix A :

$$A = \begin{pmatrix} -0.0130 & 9.5816 & 1.7946 & 0.5159 & 8.0889 & 5.6739 \\ 9.4743 & -3.4951 & 3.1129 & 0.6151 & 0.3904 & 4.8623 \\ -8.5631 & -4.6841 & 5.0933 & -3.1419 & 5.2874 & -0.6342 \\ -0.9186 & 0.0145 & -1.1929 & -9.3992 & 1.9792 & 4.7212 \\ -0.6706 & -2.9496 & 10.1276 & 4.2102 & -1.9950 & 2.7114 \\ 5.3417 & -1.0855 & 4.0736 & -0.4825 & 6.2722 & -6.1910 \end{pmatrix}. \quad (11)$$

We can obtain the eigenvector corresponding to the maximum sum of real part and imaginary part of eigenvalues of A , as follows:

$$z(t) = \begin{pmatrix} -0.3661 + 0.3696i \\ -0.2283 + 0.2939i \\ -0.3082 - 0.4881i \\ -0.0526 + 0.0305i \\ -0.3732 - 0.1860i \\ -0.2772 + 0.0571i \end{pmatrix} = (-0.5340 - 0.8456i) \begin{pmatrix} -0.1170 - 0.5069i \\ -0.1266 - 0.3499i \\ 0.5773 + 0.0000i \\ 0.0023 - 0.0607i \\ 0.3565 - 0.2162i \\ 0.0998 - 0.2648i \end{pmatrix} \quad (12)$$

$$= (-0.5340 - 0.8456i)v_{c1}$$

From the above, the eigenvector that we get is constant multiple of the eigenvector obtained from the direct calculation v_{c1} . The corresponding eigenvalue is $\lambda = \frac{\overline{z(t)}^T Az(t)}{\overline{z(t)}^T z(t)} = 11.0000 + 9.0000i$, which is just the largest sum of real part and imaginary part eigenvalues of A .

Fig. 1 illustrates the dynamic behavior of the modulus of largest sum of real part and imaginary part eigenvalues of A , and Fig. 2 illustrates the dynamic behavior of six components' modulus of the corresponding eigenvector. From these figures, we note that the proposed algorithm also has the fast convergence property, which is just one virtue of parallel computing. In addition, the proposed algorithm is not sensitive to initial value. This virtue makes actual operation a lot more convenient.

5. Conclusion

Based on the classical real domain neural network, this paper proposed a novel complex neural network to directly compute the largest sum of real part and imaginary part of eigenvalues and the corresponding eigenvectors of real normal matrices. Simulation experiment indicated that the proposed algorithm is effective.

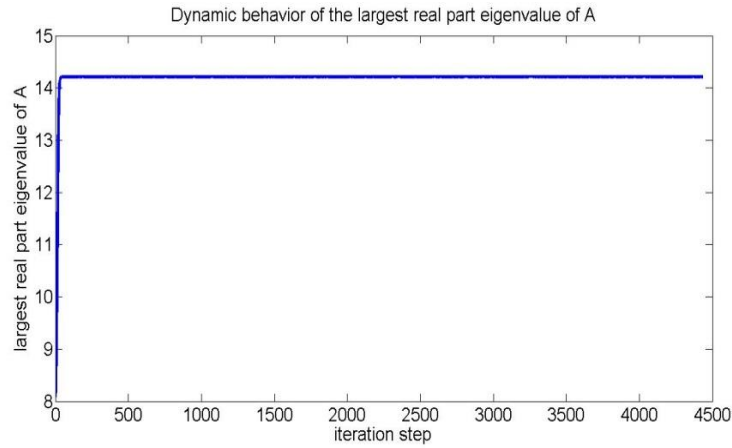


Fig. 1. Dynamic behavior of the modulus of the maximum sum of real part and imaginary part of eigenvalues of A . It should converge to $|\lambda_1| = |11 + 9i| = 14.2127$.

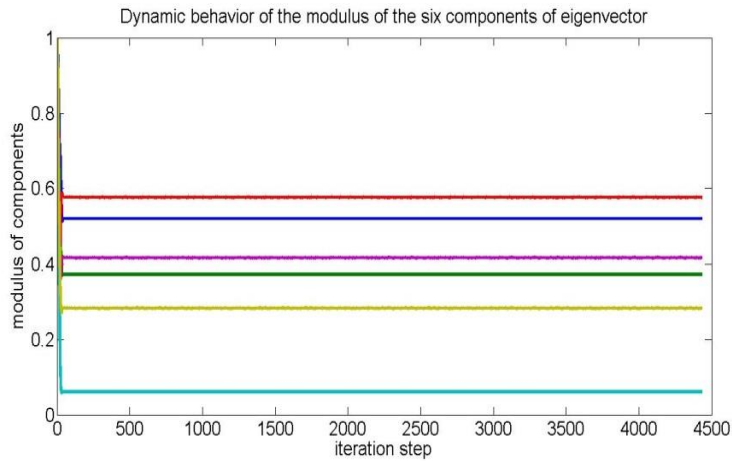


Fig. 2. Dynamic behavior of the modulus of the six components of eigenvector corresponding to the maximum sum of real parts and imaginary parts of eigenvalues. It should converge to the six components' modulus of eigenvector corresponding to eigenvalue $\lambda_1 = 11.0000 + 9.0000i$.

Acknowledgments

This work was partially supported by the General project of Sichuan Provincial Department of Education (15ZB0347, 14ZB0331), the Major project of Sichuan Provincial Department of Education (15ZA0334), The Chengdu Normal University General project (CS14ZB03) and The Chengdu Normal University

introduces the talented person scientific research start funds subsidization project (YJRC2014-5).

References

1. Oja E, A simplified neuron algorithm as a principal component analyzer, *Journal of Mathematical Biology* ,vol.15, (1982), pp.267-273.
2. F.L. Luo, R. Unbehauen and Y.D. Li, A principal component analysis algorithm with invariant norm, *Neurocomputing* ,vol.8, (1995), pp.213-221.
3. Y. Zhang, F. Yan, and H.J. Tang, Neural networks based approach for computing eigenvectors and eigenvalues of symmetric matrix, *Comput. Math. Appl*, vol.47, (2004), pp.1155-1164.
4. Y. Liu, Z.S. You and L.P. Cao, A concise functional neural network computing the largest modulus eigenvalues and their corresponding eigenvectors of a real skew matrix, *Theoretical Computer Science* ,vol.367, (2006), pp. 273-285.
5. Yiguang Liu, Zhisheng You and Liping Cao, A Recurrent Neural Network Computing the Largest Imaginary or Real Part of Eigenvalues of Real Matrices, *Computers and Mathematics with Applications*, vol.53, (2007), pp.41-53.
6. Ying Tang and J.P. Li, Another neural network based approach for computing eigenvalues and eigenvectors of real skew-symmetric matrices, *Computers and Mathematics with Applications*, vol.60, (2010), pp.1385-1392.
7. T. D. Nguye and I. Yamada, Adaptive normalized quasi-Newton algorithms for extraction of generalized eigen-pairs and their convergence analysis, *IEEE Trans. Signal Process*, vol. 61, (2013), pp.1404-1418.
8. Xiaowei Feng, Xiangyu Kong, Zhansheng Duan and Hongguang Ma, Adaptive Generalized Eigen-Pairs Extraction Algorithms and Their Convergence Analysis, *IEEE TRANSACTIONS ON SIGNAL PROCESSING*, vol. 64, (2016), pp.2976-2989.