A novel complex neural network model for computing the largest real part of eigenvalues and the corresponding eigenvector of a real matrix

Rong Ye†, Hang Tan, Xue-Song Liang and Ping-Li Wan
College of physics and engineering, Chengdu Normal University, Chengdu, 611130, China
†E-mail: yj1987211@163.com
www.cdnu.edu.cn

A novel complex neural network model was proposed, which can be used to compute the largest real part of eigenvalues and the corresponding eigenvector of a general real matrix in this work. Because of the smart regulatory factor of the model, the largest real part also can be extracted in the case of all the real parts of eigenvalues less than 0. Meanwhile, the presented paper 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 shows that the proposed model has good performance for a general real matrix.

Keywords: Complex neural network; real matrix; largest real part; eigenvalue; eigenvector

1. Introduction

In the study, we proposed a novel neural network algorithm for computing the largest real part of eigenvalues of a general real matrix, even if all the real parts are less than 0. We know that the neural network technology was first used to eigen-pairs problems about in the 1980’s [1], which is also known as principal component analysis (PCA), and then motivated broad interests from engineering and theoretical research [2-13]. In the very recent years, some adaptive generalized eigen-pairs extraction algorithms of Hermite matrices have been developed by some authors [8].

Some years ago, Liu et al. [4] proposed a more simpler 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. However, if the largest real part of eigenvalues less than 0, the algorithm cannot work correctly. In the present paper, we intend try to extract the eigenvalue and the corresponding eigenvector from the largest real part of eigenvalues of a general real matrix.
The rest of this manuscript is organized as follows: In section 2, we proposed the novel complex domain neural network model for our purpose, and the convergence analysis was presented in section 3. Numerical example given in section 4 and we summarized this paper in the last section.

2. The Novel Complex Neural Network Model

We consider the following neural network model, which is different from \([4,7]\):

\[
\frac{dv(t)}{dt} = Av(t) - (|v(t)|^2 - u)v(t).
\] (1)

Where \(v(t) \in \mathbb{R}^n\) is the \(n\)-dimension real column vectors which denotes the states of neurons, and \(u\) is a regulatory factor which can be regulated according to the actual conditions. In this paper, we intend to extract the largest real part of eigenvalue of any real matrix.

In Eq. (1), the matrix \(A\) was replaced with \(A'\):

\[
A' = \begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix},
\]

Where \(A\) is a general real matrix, \(v(t) \in \mathbb{R}^{2n}\), and with the following assumptions \(v(t)^T = [x(t)^T, y(t)^T]\), \(z(t) = x(t) + iy(t)\). Consequently, Eq. (1) turns into as follows:

\[
\frac{dz(t)}{dt} = Az(t) - [z^T(t)\bar{z}(t) - u]z(t),
\] (2)

Eq. (2) is the complex neural network model in our study, where \(z(t) \in \mathbb{C}^n\). The algorithm can be used to extract the eigenvector and the corresponding eigenvalue, which has largest real part.

All of the eigenvalues of \(A\) are denoted as \(\lambda_k^R + \lambda_k^R i, \lambda_k^R + \lambda_k^R i, \ldots, \lambda_k^R + \lambda_k^R i\). \((\lambda_k^R, \lambda_k^R) \in \mathbb{R}, k = 1, 2, \ldots, n\), and the corresponding normalized complex eigenvectors are denoted as \(S_1, S_2, \ldots, S_n\). Obviously, they construct a normalized basis in \(\mathbb{C}^{nn}\). Assume that

\[
z(t) = \sum_{k=1}^{n} z_k(t)S_k = \sum_{k=1}^{n} [x_k(t) + iy_k(t)]S_k, \] (3)
Where $z_k(t) = x_k(t) + iy_k(t)$ refers to 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 $S_k$, and the analytic expression of $|z(t)|^2$ can be written as

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

(4)

**Proof.** Assume that all eigenvalues of $A$ can be denoted as $\lambda_k^R + i\lambda_k^I$, $k = 1, \ldots, n$, where $\lambda_k^R$ and $\lambda_k^I$ are the real part and the imaginary part of the eigenvalue $\lambda_k$, respectively. Substituting Eq. (3) into Eq. (2) with $A S_k = \Lambda_k S_k$.

$$\sum_{k=1}^{n} \frac{dx_k(t)}{dt} + i \frac{dy_k(t)}{dt} S_k = \sum_{k=1}^{n} \lambda_k (x_k + iy_k) S_k - [\sum_{j=1}^{n} |z_j(t)|^2 - u] \sum_{k=1}^{n} (x_k + iy_k) S_k.$$  

(5)

Along with $S_k$, and we plug $\lambda_j = \lambda_j^R + i\lambda_j^I$, $\overline{\lambda_j} = \lambda_j^R - i\lambda_j^I$, $\lambda_k = \lambda_k^R + i\lambda_k^I$ into Eq. (5). After separating the real part and the imaginary part, and considering that $\frac{d|z_k(t)|^2}{dt} = 2x_k \frac{dx_k}{dt} + 2y_k \frac{dy_k}{dt}$, one can obtain the following equation:

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

(6)

From Eq. (6), if $|z_k(t)|^2 \neq 0$, $|z_j(t)|^2 \neq 0$, then

$$\frac{1}{|z_k(t)|^2} \frac{d}{dt} |z_k(t)|^2 - \frac{1}{|z_j(t)|^2} \frac{d}{dt} |z_j(t)|^2 = 2[\lambda_k^R - \lambda_j^R].$$  

(7)

Therefore

$$\frac{|z_k|^2}{|z_j|^2} = \frac{|z_k(0)|}{|z_j(0)|^2} \exp[2(\lambda_k^R - \lambda_j^R)t].$$  

(8)

From Eq. (6), we can directly obtain following form:
\[
\frac{d}{dt} \left| z_k(t) \right|^2 + 2(\lambda_k^R + u) \frac{1}{|z_k(t)|^2} = 2 \sum_{j=1}^{n} \left| z_j(t) \right|^2.
\]

One gets the following Eq. (10) from the above Eq. (8) and Eq. (9):

\[
\frac{d}{dt} \exp[2(\lambda_k^R + u)t] \frac{1}{|z_k(t)|^2} = 2 \sum_{j=1}^{n} \left| z_j(0) \right|^2 \exp[2(\lambda_k^R + u)t],
\]

Furthermore, with the integral calculations on both sides of Eq. (10) from 0 to \( t \), we directly get result:

\[
\left| z_k(t) \right|^2 = \exp[2(\lambda_k^R + u)t] \frac{1}{|z_k(0)|^2} \int_0^t \exp[2(\lambda_k^R + u)\tau]d\tau.
\]

Therefore one gets Eq. (4). □

### 3. Convergence Analysis

If an equilibrium vector of Eq. (2) exists, then let \( \xi \) denote it:

\[
\xi = \lim_{t \to \infty} z(t),
\]

**Theorem 3.1.** If \( A \) is not a zero matrix, and \( u = 0 \), then \( \|\xi\| \neq 0 \).

**Proof.** Obviously, from the theorem 1, we have

\[
\left| \xi \right| = \lim_{t \to \infty} \left| z(t) \right| = \lim_{t \to \infty} \sqrt{\frac{\exp[2(\lambda_k^R + u)t] \left| z_k(0) \right|^2}{1 + 2 \sum_{j=1}^{n} \left| z_j(0) \right|^2 \int_0^t \exp[2(\lambda_j^R + u)\tau]d\tau}}.
\]

\( \left| \xi \right| = 0 \) implies that all eigenvalues of \( A \) are 0, so \( A \) should be a zero matrix. □

**Theorem 3.2.** If each eigenvalue of \( A \) satisfies that \( \lambda_k^R < 0 \) (\( k = 1, 2, \ldots, n \)), and \( u \leq 0 \), then \( \left| \xi \right| = 0 \).

**Proof.** From theorem 2.1 and theorem 3.1, we easily obtain that

\[
\sum_{k=1}^{n} \exp[2(\lambda_k^R + u)t] \left| z_k(0) \right|^2 = 0,
\]

so \( \|\xi\| = 0 \). In this case, the model Eq. (3) can’t work correctly, but if we regulate the factor \( u \) satisfying \( \lambda_k^R + u > 0 \), then the model will work again. □

**Theorem 3.3.** Denoting \( (\lambda_n^R + u) = \max_{1 \leq k \leq n} (\lambda_k^R + u) \), if \( (\lambda_n^R + u) \geq 0 \), then \( \xi^T \xi = (\lambda_n^R + u) \).

**Proof.** From Eq. (11) and theorem 1, we have
Theorem 3.4. For Eq. (2), \( \lim_{t \to \infty} \frac{\xi(t)A_z(t)}{\xi^T(t)z(t)} \) will converge to the eigenvalue which has the largest real part, and the corresponding eigenvector is \( \xi = \lim_{t \to \infty} z(t) \).

Proof. From the theorem 2.1 and theorem 3.3, one can easily obtain that

\[
\lim_{t \to \infty} \frac{\xi^T(t)Az(t)}{\xi^T(t)z(t)} = \lambda_m.
\]

With \( \lim_{t \to \infty} \xi(t)A_z(t) = \lambda_m \lim_{t \to \infty} \xi^T(t)z(t) \), \( \lim_{t \to \infty} \xi^T(t)[Az(t) - \lambda_m z(t)] = 0 \), from theorem 3.1, \( \lim_{t \to \infty} \xi^T(t) \neq 0 \), so we have \( \lim_{t \to \infty}[Az(t) - \lambda_m z(t)] = 0 \). Directly, we can get \( A \lim_{t \to \infty} z(t) = \lambda_m \lim_{t \to \infty} z(t) \). So \( \xi = \lim_{t \to \infty} z(t) \) is the eigenvector corresponding to the eigenvalue \( \lambda_m \), which has the largest real part of eigenvalues.

4. Simulation Results

Example. Considering the following 6 \( \times \) 6 real artificial matrix \( A \):

\[
A = \begin{bmatrix}
-1.4624 & -9.2267 & -3.6606 & 0.2360 & -1.1129 & 2.1761 \\
-1.8325 & 2.4904 & -9.4010 & 2.4515 & -0.6820 & -5.3409 \\
0.6165 & -1.6230 & 0.9822 & -8.1747 & -2.0479 & 4.6292 \\
-3.5625 & -0.6611 & 5.1298 & -3.9747 & 3.3991 & -10.4066 \\
\end{bmatrix}
\]
The six eigenvalues of $A$ are $-11 \pm 9i, -10 \pm 4i, -8, -6$, respectively. We use $u = 10$ and the following stochastic initial value $z(0)$ for running the complex neural network algorithm Eq.(3) to obtain the largest real part of eigenvalues of matrix $A$:

$$z(0) = (1,1,1,1,1,1)^T + i(1,1,1,1,1,1)^T$$

(12)

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

$$z(t) = \begin{pmatrix} -0.2365 - 0.2318i \\ -0.5444 - 0.5336i \\ 0.7219 + 0.7076i \\ 0.9007 + 0.8829i \\ -0.5781 - 0.5667i \\ -0.1451 - 0.1423i \end{pmatrix} = (1.4282 + 1.4001i)$$

(13)

From the above, the eigenvector that we get is constant multiple of the eigenvector obtained from the direct calculation values $v_{i1}$. The corresponding eigenvalue is

$$\lambda = \frac{z(t)^T A(t)}{z(t)^T z(t)} = -6.0000$$

which is just the largest real part of eigenvalues of $A$.

Fig.1 illustrates the dynamic behavior of the largest real part of eigenvalues of $A$, and Fig.2 illustrates the dynamic behavior of six components’ modulus of the corresponding eigenvector. Fig.3 depicts the dynamic behavior of the $\xi(t)^T \xi(t)$, which should be converge to $\lambda^R + u = -6 + 10 = 4$. From these figures, we note that the proposed algorithm also has the fast convergence property, which is one advantage of parallel computing.
Fig. 1. Dynamic behavior of the maximum real part of eigenvalues of $A$. It should converge to $\lambda_0 = -6.0000$.

Fig. 2. Dynamic behavior of the modulus of the six components of eigenvector corresponding to the maximum real part eigenvalue. It should converge to the six components’ modulus of eigenvector corresponding to eigenvalue $\lambda_0 = -6.0000$. 
Fig. 3. Dynamic behavior of $\bar{\xi}(t)^T \bar{Z}(t)$, which should converge to $\lambda^*_M + u = -6 + 10 = 4$.

5. Conclusion

Based on the novel real domain neural network, this paper proposed a novel complex neural network to compute the largest real part of eigenvalues and the corresponding eigenvectors of any real matrix. Because of the model has a smart regulatory factor, the largest real part also can be extracted in the case of all the real parts of eigenvalues less than 0. We also provide 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 shows that the proposed model has good performance for a general real matrix.

Acknowledgments

This work was supported by the Major project of Education Department of Sichuan Province (15ZA0334), and partially supported by the General project of Education Department of Sichuan Province (14ZB0331), and the Introducing talents project of Chengdu Normal University (YJRC2014-5).

References


