Research Article

Neural-Network-Based Approach for Extracting Eigenvectors and Eigenvalues of Real Normal Matrices and Some Extension to Real Matrices

Xiongfei Zou,[1] Ying Tang,[2] Shirong Bu,[1] Zhengxiang Luo,[1] and Shouming Zhong[3]

[1] *School of Optoelectronic Information, University of Electronic Science and Technology of China, Chengdu 610054, China*
[2] *School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China*
[3] *School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu 611731, China*

Correspondence should be addressed to Xiongfei Zou; zouxiongfei@sohu.com

This paper introduces a novel neural-network-based approach for extracting some eigenpairs of real normal matrices of order n. Based on the proposed algorithm, the eigenvalues that have the largest and smallest modulus, real parts, or absolute values of imaginary parts can be extracted, respectively, as well as the corresponding eigenvectors. Although the ordinary differential equation on which our proposed algorithm is built is only n-dimensional, it can succeed to extract n-dimensional complex eigenvectors that are indeed $2n$-dimensional real vectors. Moreover, we show that extracting eigen-pairs of general real matrices can be reduced to those of real normal matrices by employing the norm-reducing skill. Numerical experiments verified the computational capability of the proposed algorithm.

1. Introduction

The problem of extracting special eigenpairs of real matrices has attracted much attention both in theory [1–4] and in many engineering fields such as real-time signal processing [5–8] and principal or minor component analysis [9–12]. For example, we may wish to get eigenvectors and the corresponding eigenvalues that (1) have the largest or smallest modulus; (2) have the largest or smallest real parts; (3) have the largest or smallest imaginary parts in absolute value. Two most popular methods for this problem are the power method and the Rayleigh quotient method in their direct forms or in the context of inverse iteration [13]. Recently, many neural-network-based methods have also been proposed to solve this problem [14–23]. However, most of those neural network based methods focused on computing eigenpairs of real symmetric matrices. The following two ordinary differential equations (ODEs):

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = Ax(t) - x(t)^T Ax(t) x(t), \tag{1}$$

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = x(t)^T x(t) Ax(t) - x(t)^T Ax(t) x(t) \tag{2}$$

were proposed by [19, 23], respectively, where A is a real symmetric matrix. Both (1) and (2) are efficient to compute the largest eigenvalue of A, as well as the corresponding eigenvector. In addition, they can succeed to compute the smallest eigenvalue of A and the corresponding eigenvector by simply replacing A with $-A$, for example,

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = -x(t)^T x(t) Ax(t) + x(t)^T Ax(t) x(t). \tag{3}$$

The following ODE for solving the generalized eigenvalue problem was proposed by [22]

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = Ax(t) - f(x(t)) Bx(t), \tag{4}$$

where A and B are two real symmetric matrices, and f can be a general form to some degree. Particularly, if B is the identity matrix, (4) can be used to solve the standard eigenvalue problem as (1) and (2).

References [16–18] extended those neural network based approaches to the case of real antisymmetric or special real matrices of order n, where the proposed neural networks can

be summarized by $2n$-dimensional ODEs since eigenvectors in those cases may be n-dimensional complex vectors, that is, $2n$-dimensional real vectors.

In this paper, we propose an approach for extracting six types of eigenvalues of n-by-n real normal matrices and the corresponding eigenvectors based on (2) or (3). Although eigenvectors of real normal matrices may be n-dimensional complex vectors, the computation of our proposed method can be achieved in n-dimensional real vector space, which can reduce the scale of networks a lot. Then, we show that any real matrix can be made arbitrarily close to a normal matrix by a series of similarity transformations, based on which our proposed algorithm can be extended to the case of arbitrary real matrices.

2. Main Results

Let $i = \sqrt{-1}$ be the imaginary unit, $\overline{x}$ the conjugate of x, and $\text{diag}[A_1, \ldots, A_p]$ a block diagonal matrix, where A_i, $i = 1, \ldots, p$, is a square matrix at the ith diagonal block. Unless specially stated, A is a real normal matrix in this paper.

Lemma 1. *In [13], A is a real normal matrix of order n if and only if there exists an orthogonal matrix U such that*

$$U^T A U = \text{diag}\left[\underbrace{\lambda_1, \ldots, \lambda_r}_{r}, \underbrace{A_{r+1}, \ldots, A_{r+k}}_{k}\right], \tag{5}$$

$$r + 2k = n, \quad 0 \leq r \leq n,$$

where

$$U = \left[\underbrace{u_1, \ldots, u_r}_{r \text{ real eigenvectors}}, \underbrace{u_{r+1}^R, u_{r+1}^I, \ldots, u_{r+k}^R, u_{r+k}^I}_{k \text{ pairs of complex eigenvectors}}\right], \tag{6}$$

$$A_{r+j} = \begin{pmatrix} a_{r+j} & |b_{r+j}| \\ -|b_{r+j}| & a_{r+j} \end{pmatrix}, \quad j = 1, \ldots, k.$$

Here, λ_j, $j = 1, \ldots, r$, are r real eigenvalues of A corresponding to the r real eigenvectors u_j, and $a_{r+s} \pm i b_{r+s}$, $s = 1, \ldots, k$, are k pairs of complex eigenvalues of A corresponding to the k pairs of complex eigenvectors $u_{r+s}^R \pm i u_{r+s}^I$.

For simplicity, let $a_j = \lambda_j$, $b_j = 0$ for $j = 1, \ldots, r$. Based on (5) and (6), it is straightforward to verify

$$U^T \left(A - A^T\right)^T \left(A - A^T\right) U$$
$$= 4 \, \text{diag}\left[b_1^2, \ldots, b_r^2, b_{r+1}^2, b_{r+1}^2, \ldots, b_{r+k}^2, b_{r+k}^2\right] \tag{7}$$

$$U^T \left(A + A^T\right) U$$
$$= 2 \, \text{diag}\left[a_1, \ldots, a_r, a_{r+1}, a_{r+1}, \ldots, a_{r+k}, a_{r+k}\right] \tag{8}$$

$$U^T \left(A A^T\right) U$$
$$= \text{diag}\left[a_1^2, \ldots, a_r^2, a_{r+1}^2 + b_{r+1}^2, a_{r+1}^2 \right. \tag{9}$$
$$\left. + b_{r+1}^2, \ldots, a_{r+k}^2 + b_{r+k}^2, a_{r+k}^2 + b_{r+k}^2\right].$$

Then, the following six definitions and two lemmas are presented, which will be involved much in the sequel.

(C1) Let $\mathscr{J}_1 = \{s \mid |b_s| = \max\{|b_1|, \ldots, |b_{r+k}|\}\}$. Then for all $j, k \in \mathscr{J}_1$, $a_j = a_k$.

(C2) Let $\mathscr{J}_2 = \{s \mid |b_s| = \min\{|b_1|, \ldots, |b_{r+k}|\}\}$. Then for all $j, k \in \mathscr{J}_2$, $a_j = a_k$.

(C3) Let $\mathscr{J}_3 = \{s \mid a_s = \max\{a_1, \ldots, a_{r+k}\}\}$. Then for all $j, k \in \mathscr{J}_3$, $|b_j| = |b_k|$.

(C4) Let $\mathscr{J}_4 = \{s \mid a_s = \min\{a_1, \ldots, a_{r+k}\}\}$. Then for all $j, k \in \mathscr{J}_4$, $|b_j| = |b_k|$.

(C5) Let $\mathscr{J}_5 = \{s \mid a_s^2 + b_s^2 = \max\{a_1^2 + b_1^2, \ldots, a_{r+k}^2 + b_{r+k}^2\}\}$. Then for all $j, k \in \mathscr{J}_5$, $a_j = a_k$.

(C6) Let $\mathscr{J}_6 = \{s \mid a_s^2 + b_s^2 = \min\{a_1^2 + b_1^2, \ldots, a_{r+k}^2 + b_{r+k}^2\}\}$. Then for all $j, k \in \mathscr{J}_6$, $a_j = a_k$.

Lemma 2 (Theorem 4 in [23]). *Assume that nonzero $x(0) \in R^n$ is not orthogonal to the eigensubspace corresponding to the largest eigenvalue of A. Then, the solution of (2) starting from $x(0)$ converges to an eigenvector corresponding to the largest eigenvalue of A that is equal to $\lim_{t \to +\infty}(x(t)^T A x(t)/x(t)^T x(t))$.*

Lemma 3 (Theorem 5 in [23]). *Assume that nonzero $x(0) \in R^n$ is not orthogonal to the eigensubspace corresponding to the smallest eigenvalue of A. Then, the solution of (3) starting from $x(0)$ converges to an eigenvector corresponding to the smallest eigenvalue of A that is equal to $\lim_{t \to +\infty}(x(t)^T A x(t)/x(t)^T x(t))$.*

Remark 4. If we randomly choose $x(0)$, the projection of $x(0)$ on the eigensubspace corresponding to the largest or smallest eigenvalue of A will be nonzero with high probability. Hence, (2) and (3) can almost work well with randomly generated $x(0)$.

2.1. Computing the Eigenvalues with the Largest or Smallest Imaginary Parts in Absolute Value, as well as the Corresponding Eigenvectors. Without loss of generality, in this subsection we assume that $b_1 = \cdots = b_r = 0 < b_{r+1}^2 \leq \cdots \leq b_{r+k}^2$ in (7). Note that $(A - A^T)^T(A - A^T) = -(A - A^T)^2$.

Based on (7), we know that 0 (if any) is the eigenvalue of $-(A - A^T)^2$ corresponding to the eigenvector u_j, $j = 1, \ldots, r$, and that $4b_{r+s}^2$, $s = 1, \ldots, k$, is the eigenvalue of $-(A - A^T)^2$ corresponding to the eigenvectors u_{r+s}^R or u_{r+s}^I. If replacing A with the symmetric matrix $-(A - A^T)^2$ in (2) as

$$\frac{dx(t)}{dt} = -x(t)^T x(t) \left(A - A^T\right)^2 x(t) + x(t)^T \left(A - A^T\right)^2 x(t) \, x(t) \tag{10}$$

based on Lemma 2, we know that $\alpha = \lim_{t \to \infty} x(t)$ is an eigenvector of $-(A - A^T)^2$ corresponding to $4b_{r+k}^2$ and $4b_{r+k}^2 = \lim_{t \to \infty}(-x(t)^T(A - A^T)^2 x(t)/x(t)^T x(t))$ (thus, $|b_{r+k}|$ has been gotten), where $x(t)$ is a solution of (10). If $b_{r+k} = 0$, b_j's

are all zero, meaning that A is symmetric. In this case, we can directly use (2) or (3) to get the largest or smallest eigenvalue of A and the corresponding eigenvector, respectively. Hence, we assume $b_{r+k} \neq 0$.

The following lemma introduces an approach for computing a_{r+k} under the condition (C1).

Lemma 5. *Assume that (C1) holds. Let $\alpha = \lim_{t \to \infty} x(t)$, where $x(t)$ is a solution of (10). Then, $a_{r+k} = \alpha^T(A + A^T)\alpha / 2\alpha^T \alpha$. In addition, there exists $a_{r+k}^2 + b_{r+k}^2 = \alpha^T(AA^T)\alpha / \alpha^T \alpha$.*

Proof. Assume that $4b_{r+j}^2$, $j = s, s+1, \ldots, k$, $s \geq 1$, are the largest eigenvalues of $-(A - A^T)^2$, that is; $|b_{r+s}| = |b_{r+s+1}| = \cdots = |b_{r+k}|$. Therefore,

$$\mathcal{J}_1 = \{r + s, \ldots, r + k\}. \tag{11}$$

Based on Lemma 2 and (7), we know that α should be a linear combination of u_j^R and u_j^I, $j \in \mathcal{J}_1$. Let

$$\alpha = \sum_{j \in \mathcal{J}_1} \left(\gamma_j u_j^R + \tau_j u_j^I \right), \quad \gamma_j, \tau_j \in R. \tag{12}$$

In addition, by (8) we have

$$\left(A + A^T \right) u_j^R = 2a_j u_j^R, \quad \left(A + A^T \right) u_j^I = 2a_j u_j^I, \quad j \in \mathcal{J}_1. \tag{13}$$

And by (9), we have

$$\left(AA^T \right) u_j^R = \left(a_j^2 + b_j^2 \right) u_j^R,$$
$$\left(AA^T \right) u_j^I = \left(a_j^2 + b_j^2 \right) u_j^I, \quad j \in \mathcal{J}_1. \tag{14}$$

Because U is an orthogonal matrix and for all $j \in \mathcal{J}_1$, $a_j = a_{r+k}$ holds due to (C1), it is straightforward to verify

$$\frac{\alpha^T \left(A + A^T \right) \alpha}{2\alpha^T \alpha} = \frac{2a_{r+k} \sum_{j \in \mathcal{J}_1} \left(\gamma_j^2 + \tau_j^2 \right)}{2 \sum_{j \in \mathcal{J}_1} \left(\gamma_j^2 + \tau_j^2 \right)} = a_{r+k},$$

$$\frac{\alpha^T \left(AA^T \right) \alpha}{\alpha^T \alpha} = \frac{\left(a_{r+k}^2 + b_{r+k}^2 \right) \sum_{j \in \mathcal{J}_1} \left(\gamma_j^2 + \tau_j^2 \right)}{\sum_{j \in \mathcal{J}_1} \left(\gamma_j^2 + \tau_j^2 \right)} = a_{r+k}^2 + b_{r+k}^2, \tag{15}$$

thus proving the lemma. $\square$

Remark 6. Based on Lemma 5, if $a_{r+k}^2 + b_{r+k}^2 \neq \alpha^T(AA^T)\alpha / \alpha^T \alpha$, (C1) does not hold surely, which can be used to check whether (C1) holds or not.

The following lemma introduces an approach for computing a pair of conjugated eigenvectors of A corresponding to the eigenvalues $a_{r+k} \pm i|b_{r+k}|$ under the condition (C1).

Lemma 7. *Assume that (C1) holds. Given any nonzero $4b_{r+k}^2$, the largest eigenvalue of $-(A - A^T)^2$, and the corresponding eigenvector α obtained by (10), let $a_{r+k} = \alpha^T(A + A^T)\alpha / 2\alpha^T \alpha$ and $\beta = (A - A^T)\alpha / 2|b_{r+k}|$. Then, $a_{r+k} \pm i|b_{r+k}|$ are two eigenvalues of A corresponding to the eigenvectors $\beta \pm i\alpha$, respectively.*

Proof. Let $\mathcal{J}_1$ and α take the form as (11) and (12), respectively. Based on (13), we can write

$$\frac{\left(A + A^T \right)}{2} \alpha = a_{r+k} \alpha. \tag{16}$$

Following the decomposition of A as (5) and the definition of β, we have

$$|b_{r+k}| \beta = \left(A - A^T \right) \frac{\alpha}{2}$$
$$= U^T \operatorname{diag} \left[\underbrace{0, \ldots, 0}_{r}, B_{r+1}, \ldots, B_{r+k} \right] U \frac{\alpha}{2}, \tag{17}$$

where

$$B_j = \begin{cases} \begin{pmatrix} 0 & 2|b_j| \\ -2|b_j| & 0 \end{pmatrix}, & j \in \{r+1, \ldots, r+k\} \setminus \mathcal{J}_1 \\ \begin{pmatrix} 0 & 2|b_{r+k}| \\ -2|b_{r+k}| & 0 \end{pmatrix}, & j \in \mathcal{J}_1. \end{cases} \tag{18}$$

Taking (12) into (17), we get

$$\beta = \sum_{j \in \mathcal{J}_1} \left(\tau_j u_j^R - \gamma_j u_j^I \right). \tag{19}$$

Based on (13) and (19), it is straightforward to verify

$$\frac{\left(A + A^T \right)}{2} \beta = a_{r+k} \beta. \tag{20}$$

In addition, since $4b_{r+k}^2$ is the eigenvalue of $-(A - A^T)^2$ corresponding to the eigenvector α, we have $-(A - A^T)^2\alpha = 4b_{r+k}^2\alpha$. Hence, $-(A - A^T)|b_{r+k}|\beta = 2b_{r+k}^2\alpha$. Since $b_{r+k} \neq 0$, we have

$$\frac{\left(A - A^T \right)}{2} \beta = -|b_{r+k}| \alpha. \tag{21}$$

By (20) and (21), we have

$$A\beta = \frac{A + A^T}{2}\beta + \frac{A - A^T}{2}\beta = a_{r+k}\beta - |b_{r+k}|\alpha. \tag{22}$$

By (16) and (17), we have

$$A\alpha = \frac{A + A^T}{2}\alpha + \frac{A - A^T}{2}\alpha = a_{r+k}\alpha + |b_{r+k}|\beta. \tag{23}$$

Then, it is straightforward to verify

$$A\left(\beta + i\alpha \right) = \left(a_{r+k} + i|b_{r+k}| \right)\left(\beta + i\alpha \right),$$
$$A\left(\beta - i\alpha \right) = \left(a_{r+k} - i|b_{r+k}| \right)\left(\beta - i\alpha \right), \tag{24}$$

thus proving the lemma. $\square$

To get $|b_1|$, we can in advance get $4b_1^2$, the smallest eigenvalue of $-(A - A^T)^2$ by replacing A with $-(A - A^T)^2$ in (3) as follows:

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = x(t)^T x(t)\left(A - A^T\right)^2 x(t) - x(t)^T\left(A - A^T\right)^2 x(t) x(t). \tag{25}$$

Let $\alpha^* = \lim_{t \to \infty} x(t)$, where $x(t)$ is a solution of (25). From Lemma 3, we know $4b_1^2 = -(\alpha^*)^T (A - A^T)^2 \alpha^*/(\alpha^*)^T \alpha^*$. Then, the following lemma similar to Lemma 5 can be used to compute a_1.

Lemma 8. *Assume that (C2) holds. Then, $a_1 = (\alpha^*)^T(A + A^T)\alpha^*/2(\alpha^*)^T\alpha^*$. In addition, there exists $a_1^2 + b_1^2 = (\alpha^*)^T(AA^T)\alpha^*/(\alpha^*)^T\alpha^*$.*

Proof. The proof is almost the same to that in Lemma 5. □

Note that b_1 may be zero; that is, $-(A - A^T)^2$ has real eigenvalues. In this case, we have the following lemma.

Lemma 9. *Assume that (C2) holds and 0 is the smallest eigenvalue of $-(A - A^T)^2$ corresponding to the eigenvector α^*. Then, a_1 is the eigenvalue of A corresponding to the eigenvector α^*, where $a_1 = (\alpha^*)^T(A + A^T)\alpha^*/2(\alpha^*)^T\alpha^*$.*

Proof. Following the conditions, we have $-(A - A^T)^2\alpha^* = 0$. Hence, $(\alpha^*)^T(A-A^T)^T(A-A^T)\alpha^* = \|(A - A^T)\alpha^*\|_2^2 = 0$; that is,

$$\left(A - A^T\right)\alpha^* = 0. \tag{26}$$

Note that $b_1 = \cdots = b_r = 0$, $r \geq 1$, because 0 is the smallest eigenvalue of $-(A - A^T)^2$. Based on the definition of $\mathscr{J}_2$, we have

$$a_1 = \cdots = a_r, \quad \mathscr{J}_2 = \{1, \ldots, r\}, \quad r \geq 1. \tag{27}$$

Applying Lemma 3 to (25), we know that α^* should be a linear combination of $u_1, \ldots, u_r$. Let

$$\alpha^* = \sum_{j \in \mathscr{J}_2} \gamma_j u_j, \quad \gamma_j \in R. \tag{28}$$

By (8), we have

$$\left(A + A^T\right)u_j = 2a_j u_j = 2a_1 u_j, \quad j \in \mathscr{J}_2. \tag{29}$$

Therefore,

$$\frac{A + A^T}{2}\alpha^* = \frac{1}{2}\left(\sum_{j \in \mathscr{J}_2} \gamma_j\left(A + A^T\right)u_j\right) = a_1\alpha^*. \tag{30}$$

Then, by (26) and (30), it is straightforward to verify

$$A\alpha^* = \frac{A + A^T}{2}\alpha^* + \frac{A - A^T}{2}\alpha^* = a_1\alpha^*, \tag{31}$$

thus proving the lemma. □

In the case of $b_1 \neq 0$, that is, all the eigenvalues of $-(A - A^T)^2$ are complex numbers, we have the following lemma similar to Lemma 7.

Lemma 10. *Assume that (C2) holds. Given any nonzero $4b_1^2$, the smallest eigenvalue of $-(A - A^T)^2$, and the corresponding eigenvector α^* obtained by (25), let $a_1 = (\alpha^*)^T(A + A^T)\alpha^*/2(\alpha^*)^T\alpha^*$ and $\beta^* = (A - A^T)\alpha^*/2|b_1|$. Then, $a_1 \pm i|b_1|$ are two eigenvalues of A corresponding to the eigenvectors $\beta^* \pm i\alpha^*$, respectively.*

Proof. The proof is almost the same to that in Lemma 7. □

Remark 11. Among the following four real normal matrices

$$\begin{pmatrix} 7 & 0 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 6 & 2 \\ 0 & 0 & 0 & -2 & 6 \end{pmatrix}, \quad \begin{pmatrix} 7 & 0 & 0 & 0 & 0 \\ 0 & 5 & 1 & 0 & 0 \\ 0 & -1 & 5 & 0 & 0 \\ 0 & 0 & 0 & 6 & 2 \\ 0 & 0 & 0 & -2 & 6 \end{pmatrix},$$
$$\begin{pmatrix} 7 & 0 & 0 & 0 & 0 \\ 0 & 5 & -2 & 0 & 0 \\ 0 & 2 & 5 & 0 & 0 \\ 0 & 0 & 0 & 5 & 2 \\ 0 & 0 & 0 & -2 & 5 \end{pmatrix}, \quad \begin{pmatrix} 7 & 0 & 0 & 0 & 0 \\ 0 & 5 & 2 & 0 & 0 \\ 0 & -2 & 5 & 0 & 0 \\ 0 & 0 & 0 & 6 & 2 \\ 0 & 0 & 0 & -2 & 6 \end{pmatrix}, \tag{32}$$

only the first three meet (C1), but the last one does not. And only the last three matrices meet (C2), but the first one does not.

2.2. Computing the Eigenvalues with the Largest or Smallest Real Parts, as well as the Corresponding Eigenvectors. As shown in (8), $2a_j$, $j = 1, \ldots, r + k$, are the eigenvalues of the symmetric matrix $(A + A^T)$. In this subsection, we assume that $a_1 \leq \cdots \leq a_r$ and $a_{r+1} \leq \cdots \leq a_{r+k}$, which can be achieved by reordering a_j and the corresponding columns of U.

If replacing A with $(A + A^T)$ in (2), we get

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = x(t)^T x(t)\left(A + A^T\right)x(t) - x(t)^T\left(A + A^T\right)x(t) x(t). \tag{33}$$

Without loss of generality, assume that a_{r+k} is the largest real part of the eigenvalues of A (it may be a_r. However, Lemmas 12 and 13 have no difference in that case). Applying Lemma 2 to (33), we can get $2a_{r+k} = \lim_{t \to \infty}(x(t)^T(A + A^T)x(t)/x(t)^T x(t))$, the largest eigenvalue of $(A + A^T)$, and the corresponding eigenvector $\zeta = \lim_{t \to \infty} x(t)$, where $x(t)$ is a solution of (33).

The following lemma introduces an approach for computing $|b_{r+k}|$ under the condition (C3).

Lemma 12. *Assume that (C3) holds. Let $\zeta = \lim_{t \to \infty} x(t)$, where $x(t)$ is a solution of (33). Then, $|b_{r+k}| = \sqrt{(\zeta^T(AA^T)\zeta/\zeta^T\zeta) - a_{r+k}^2}$.*

Proof. By Lemma 2, we know that ζ should be a linear combination of u_j^R and u_j^I, $j \in \mathscr{J}_3$. Let

$$\zeta = \sum_{j \in \mathscr{J}_3} \left(\gamma_j u_j^R + \tau_j u_j^I \right), \quad \gamma_j, \tau_j \in R. \tag{34}$$

Then, based on (9), we get

$$\left(AA^T \right) u_j^R = \left(a_j^2 + b_j^2 \right) u_j^R,$$
$$\left(AA^T \right) u_j^I = \left(a_j^2 + b_j^2 \right) u_j^I, \quad j \in \mathscr{J}_3. \tag{35}$$

Because U is an orthogonal matrix and for all $j \in \mathscr{J}_3$, $a_j = a_{r+k}$ and $|b_j| = |b_{r+k}|$ hold, that is, $a_j^2 + b_j^2 = a_{r+k}^2 + b_{r+k}^2$ for all $j \in \mathscr{J}_3$, we have

$$\frac{\zeta^T \left(AA^T \right) \zeta}{\zeta^T \zeta} = \frac{\left(a_{r+k}^2 + b_{r+k}^2 \right) \sum_{j \in \mathscr{J}_3} \left(\gamma_j^2 + \tau_j^2 \right)}{\sum_{j \in \mathscr{J}_3} \left(\gamma_j^2 + \tau_j^2 \right)} = a_{r+k}^2 + b_{r+k}^2, \tag{36}$$

thus proving the lemma. □

The following lemma introduces an approach for computing a pair of conjugated eigenvectors of A corresponding to the eigenvalues $a_{r+k} \pm i|b_{r+k}|$ under the condition (C3).

Lemma 13. *Assume that (C3) holds. Given any nonzero $2a_{r+k}$, the largest eigenvalue of $(A + A^T)$, and the corresponding eigenvector ζ obtained by (33), let $|b_{r+k}| = \sqrt{(\zeta^T(AA^T)\zeta/\zeta^T\zeta) - a_{r+k}^2}$. If $b_{r+k} = 0$, a_{r+k} is an eigenvalue of A corresponding to the eigenvector ζ. If $b_{r+k} \neq 0$, let $\eta = (A - A^T)\zeta/2|b_{r+k}|$. Then, $a_{r+k} \pm i|b_{r+k}|$ are two eigenvalues of A corresponding to the eigenvectors $\eta \pm i\zeta$, respectively.*

Proof. Combining the proofs of Lemmas 9 and 7, we can prove this lemma. □

If replacing A with $(A + A^T)$ in (3), we get

$$\frac{dx(t)}{dt} = - x(t)^T x(t) \left(A + A^T \right) x(t) + x(t)^T \left(A + A^T \right) x(t) x(t). \tag{37}$$

Without loss of generality, assume that a_{r+1} is the smallest real part of the eigenvalues of A (it may be a_1. However, Lemma 14 has no difference in that case). Applying Lemma 3 to (37), we can obtain $2a_{r+1}$, the smallest eigenvalue of $(A+A^T)$, as well as the corresponding eigenvector, denoted by ζ^*. Then, we have the following lemma.

Lemma 14. *Assume that (C4) holds. Given any nonzero $2a_{r+1}$, the smallest eigenvalue of $(A + A^T)$, and the corresponding eigenvector ζ^* obtained by (37), let $|b_{r+1}| = \sqrt{((\zeta^*)^T(AA^T)\zeta^*/(\zeta^*)^T\zeta^*) - a_{r+1}^2}$. If $b_{r+1} = 0$, a_{r+1} is the eigenvalue of A corresponding to the eigenvector ζ^*. If $b_{r+1} \neq 0$, let $\eta^* = (A - A^T)\zeta^*/2|b_{r+1}|$. Then, $a_{r+1} \pm i|b_{r+1}|$ are two eigenvalues of A corresponding to the eigenvectors $\eta^* \pm i\zeta^*$, respectively.*

Proof. Combining the proofs of Lemmas 12, 9, and 7, we can prove this lemma. □

Remark 15. Among the following four real normal matrices

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & -2 & 1 \end{pmatrix}, \quad \begin{pmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & -2 & 1 \end{pmatrix},$$
$$\begin{pmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & 4 & 2 & 0 & 0 \\ 0 & -2 & 4 & 0 & 0 \\ 0 & 0 & 0 & 4 & 2 \\ 0 & 0 & 0 & -2 & 4 \end{pmatrix}, \quad \begin{pmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 5 & 2 \\ 0 & 0 & 0 & -2 & 5 \end{pmatrix}, \tag{38}$$

only the first three meet (C3), but the last one does not. And only the last three matrices meet (C4), but the first one does not.

2.3. Computing the Eigenvalues with the Largest or Smallest Modulus, as well as the Corresponding Eigenvectors. Reorder the eigenvalues of the symmetric matrix AA^T in (9) and the corresponding columns of U such that $a_1^2 \leq \cdots \leq a_r^2$ and $a_{r+1}^2 + b_{r+1}^2 \leq \cdots \leq a_{r+k}^2 + b_{r+k}^2$. Without loss of generality, assume that $a_{r+k} \pm i|b_{r+k}|$ are the eigenvalues of A that have the largest modulus (it may be a_r. However, Lemma 16 has no difference in that case), and that $a_{r+1} \pm i|b_{r+1}|$ are the eigenvalues of A that have the smallest modulus (it may be a_1. However, Lemma 17 has no difference in that case).

Replacing A with AA^T in (2), we get

$$\frac{dx(t)}{dt} = x(t)^T x(t) AA^T x(t) - x(t)^T AA^T x(t) x(t). \tag{39}$$

Applying Lemma 2 to (39), we can obtain $a_{r+k}^2 + b_{r+k}^2 = \lim_{t \to \infty}(x(t)^T A^T A x(t)/x(t)^T x(t))$, the largest eigenvalue of AA^T, and the corresponding eigenvector $\theta = \lim_{t \to \infty} x(t)$, where $x(t)$ is the solution of (39). Then, we have the following lemma.

Lemma 16. *Assume that (C5) holds. Given any $a_{r+k}^2 + b_{r+k}^2$, the largest eigenvalue of AA^T, and the corresponding eigenvector θ obtained by (39). Then, $a_{r+k} = \theta^T(A + A^T)\theta/2\theta^T\theta$. Thus, $|b_{r+k}|$ can be gotten. If $b_{r+k} = 0$, a_{r+k} is an eigenvalue of A corresponding to the eigenvector θ. If $b_{r+k} \neq 0$, let $\omega = (A - A^T)\theta/|2b_{r+k}|$. Then, $a_{r+k} \pm i|b_{r+k}|$ are two eigenvalues of A corresponding to the eigenvectors $\omega \pm i\theta$, respectively.*

Proof. Combining the proofs of Lemmas 5, 7, and 9, we can prove this lemma. □

Replacing A with AA^T in (3), we get

$$\frac{dx(t)}{dt} = -x(t)^T x(t) AA^T x(t) + x(t)^T AA^T x(t) x(t). \tag{40}$$

Applying Lemma 3 to (40), we can obtain $a_{r+1}^2 + b_{r+1}^2$, the smallest eigenvalue of AA^T, and the corresponding eigenvector, denoted by θ^*. Then, we have the following lemma.

Lemma 17. *Assume (C6) holds. Given any $a_{r+1}^2 + b_{r+1}^2$, the smallest eigenvalue of AA^T, and the corresponding eigenvector θ^* obtained by (40). Then, $a_{r+1} = (\theta^*)^T (A + A^T)\theta^*/2(\theta^*)^T \theta^*$. So one can get that $|b_{r+1}|$. If $b_{r+1} = 0$, a_{r+1} is an eigenvalue of A corresponding to the eigenvector θ^*. If $b_{r+1} \neq 0$, let $\omega^* = (A - A^T)\theta^*/2|b_{r+1}|$. Then, $a_{r+1} \pm i|b_{r+1}|$ are two eigenvalues of A corresponding to the eigenvectors $\omega^* \pm i\theta^*$, respectively.*

Proof. Combining the proofs of Lemmas 5, 7, and 9, we can prove this lemma. □

Remark 18. Among the following four real normal matrices

$$\begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 & 0 \\ 0 & 0 & 6 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & -2 & 1 \end{pmatrix}, \quad \begin{pmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 & 0 \\ 0 & 0 & 6 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & -2 & 1 \end{pmatrix},$$

$$\begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & -4 & 0 \\ 0 & 4 & 3 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & -4 \end{pmatrix}, \quad \begin{pmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 3 & 4 \\ 0 & 0 & 0 & -4 & 3 \end{pmatrix}, \tag{41}$$

only the first three meet (C5), but the last one does not. And only the last three meet (C6), but the first one does not.

However, there exists some specially constructed A that meet none of (C1) to (C6), for example, $A = \text{diag}[A_1, A_2, A_3, A_4]$, where

$$A_1 = \begin{pmatrix} -4 & 1 \\ -1 & -4 \end{pmatrix}, \quad A_2 = \begin{pmatrix} -4 & 3 \\ -3 & -4 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 4 & 3 \\ -3 & 4 \end{pmatrix},$$

$$A_4 = \begin{pmatrix} 4 & 1 \\ -1 & 4 \end{pmatrix}. \tag{42}$$

Nevertheless, a randomly generated real normal matrix can meet (C1) to (C6) with high probability.

2.4. Extension to Arbitrary Real Matrices. In this subsection, A is an arbitrary real matrix of order n. Let $\|A\|$ be the Frobenius norm of A, and, λ_j, $j = 1, \ldots, n$, be the eigenvalue of A. Denote the set of all complex nonsingular matrices by $\mathcal{T}$.

By the Schur inequality [13], we know

$$\sum_{j=1}^n |\lambda_j^2| \leq \|A\|^2 \tag{43}$$

with equality if and only if A is a normal matrix. Since the spectrum of A does not change by a similarity transformation, the inequality

$$\sum_{j=1}^n |\lambda_j^2| \leq \|P^{-1}AP\|^2, \quad P \in \mathcal{T} \tag{44}$$

holds with equality if and only if $P^{-1}AP$ is a normal matrix. In addition, [24] proved

$$\sum_{i=1}^n |\lambda_i^2| = \inf_{P \in \mathcal{T}} \|P^{-1}AP\|^2. \tag{45}$$

Based on (45), if we can find a sequence A_j as follows:

$$A_1 = A, \quad A_{j+1} = P_j^{-1} A_j P_j, \quad \|A_{j+1}\| < \|A_j\|,$$
$$P_j \in \mathcal{T}, \quad j = 1, 2, \ldots \tag{46}$$

such that

$$\sum_{i=1}^n |\lambda_i^2| = \lim_{j \to \infty} \|A_j\|^2, \tag{47}$$

where $A_\infty = \lim_{j \to \infty} A_j$ is to be a normal matrix with the same eigenvalues as A. Such skill, termed as the norm-reducing technique, has been proposed by [25–27]. Moreover, following the idea presented by [26], it is easy to find that when A is real, A_∞ can be chosen to be a real normal matrix.

In a word, any real matrix A can be translated into a real normal matrix $A_\infty = P_\infty^{-1} A P_\infty$ by a similarity transformation $P_\infty = P_1 P_2 \cdots$. Typical approaches for constructing P_j can be found in [26]. Note that if λ is the eigenvalue of A_∞ corresponding to the eigenvector u, λ is the eigenvalue of A corresponding to the eigenvector $P_\infty u$. Hence, our proposed algorithm can be extended to extract eigenpairs of arbitrary real matrices by employing the norm-reducing technique.
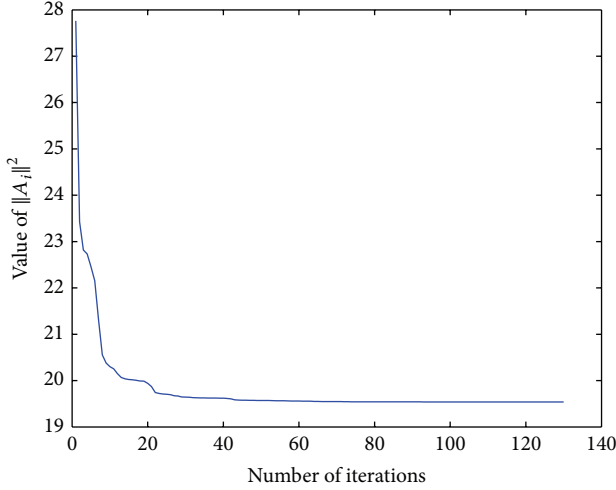
Without loss of generality, we use the following random matrix A as an example to describe the norm-reducing technique:

$$A = \begin{pmatrix} 1.0541 & 0.4052 & -1.0191 & -0.5771 & 0.4158 \\ -1.9797 & -0.7025 & -1.3852 & -0.8364 & 0.0430 \\ -1.8674 & 1.4990 & 0.9549 & 0.8530 & -0.9489 \\ -1.8324 & 0.1378 & -0.6011 & 0.4773 & 0.5416 \\ 0.8486 & -1.5868 & -1.1719 & 0.3023 & -0.8211 \end{pmatrix}. \tag{48}$$

The Frobenius norm of A is $\|A\| = \sqrt{27.7629}$, and the eigenvalues of matrix A are $\lambda_1 = 2.8797$, $\lambda_2 = \overline{\lambda}_3 = -0.6900 + 1.8770i$, $\lambda_4 = -1.5144$, and $\lambda_5 = 0.9774$; obviously $\|A\| > \sqrt{\sum_{i=1}^5 \lambda_i^2} = 19.5401$, so A is a nonnormal real matrix. According to the approach that presented in [26], we can construct P_j, $j = 1, 2, \ldots$. If the condition $\|A_i - A_{i+1}\| < 10^{-10}$ can be satisfied, we break the iteration. After 129 iterations, we can obtain a approximate normal matrix A_{130}:

$$A_{130} = \begin{pmatrix} -0.7441 & 0.0007 & 1.6334 & 0.8210 & 0.0010 \\ -0.0035 & 2.8797 & -0.0038 & 0.0000 & -0.0000 \\ -1.4162 & -0.0048 & -0.9132 & 0.8223 & 0.0034 \\ -1.1560 & -0.0019 & -0.1181 & -1.2370 & -0.0014 \\ 0.0011 & -0.0000 & 0 & -0.0036 & 0.9774 \end{pmatrix}, \tag{49}$$

from which we have $\|A_{130}^* A_{130} - A_{130} A_{130}^*\| = \sqrt{0.00121231723027107}$ and $\|A_{130}\| = \sqrt{19.540133042214}$

FIGURE 1: Trajectory of $\|A_i\|^2$.

FIGURE 2: Average performance trajectory of norm-reducing technique.

that is very close to $\sqrt{\sum_{i=1}^{5} \lambda_i^2} = \sqrt{19.5400526824864}$; so the matrix A_{130} can be regarded as a normal matrix in practical application, the corresponding $P_{\infty} = P_1 P_2 P_3 \cdots P_{129}$ are as follows:

$$
P_{\infty} = \begin{pmatrix}
-0.3499 & -0.5290 & -0.0090 & 0.2228 & -0.0151 \\
-0.4809 & -0.0259 & -0.6935 & -0.4079 & 0.1559 \\
-0.2914 & 0.7211 & 0.6743 & 0.0395 & 0.3437 \\
-0.4484 & 0.1475 & -0.3668 & 0.2444 & -0.8704 \\
0.0611 & -0.3265 & 0.4294 & -1.1028 & -0.5166
\end{pmatrix},
$$
(50)

which satisfy $A_{130} = P_{\infty}^{-1} A P_{\infty}$. We also can see that the eigenvalues of A_{130} are $\lambda_1 = -1.5144$, $\lambda_2 = \overline{\lambda_3} = -0.6900 + 1.8770i$, $\lambda_4 = 2.8797$, and $\lambda_5 = 0.9774$, which are just the eigenvalues of original nonnormal matrix A. We presented here the transient behavior of $\|A_i\|^2$ in Figure 1, where $i = 1, 2, 3, \ldots$ is the iterations and $A_1 = A$.

In the following, we use $N = 1000$ random matrices to verify the average performance of the norm-reducing technique. Let

$$
\text{Mean } \Delta_i^2 = \frac{1}{N} \sum_{j=1}^{N} \Delta_{ji}^2
$$
(51)

denotes the average measure of a large number of nonnormal matrices in a statistical sense at ith iteration, where $\Delta_{ji}^2 = \|A_{ji}^T A_{ji} - A_{ji} A_{ji}^T\|^2$ is the measure of the nonnormal matrix A_j at ith iteration, $\Delta_{ji}^2 = 0$ if and only if A_j is normal matrix at ith iteration. We also presented the dynamic behavior trajectory of Mean Δ^2 in Figure 2, from which we can see that for most of nonnormal matrices, after 350 iterations, the Mean Δ^2 is very close to zero.

3. Neural Implementation Description

In the presented paper, we mainly focus on the classical neural network differential equation as shown in (2), where
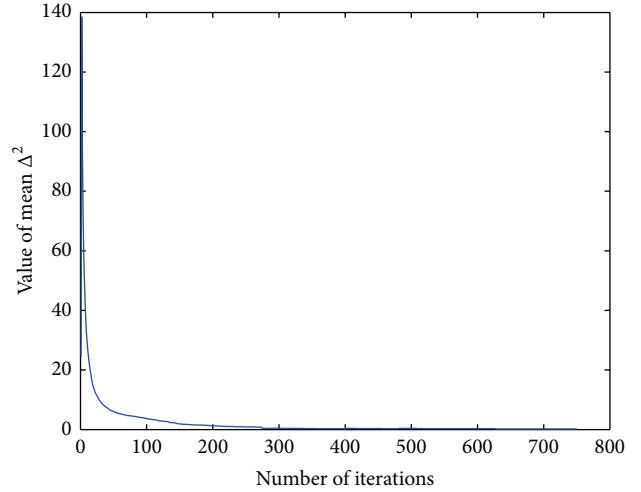
$A = (a_{ij})$, $i, j = 1, 2, \ldots, n$ are symmetric matrices that need to calculate eigenvalues and the corresponding eigenvectors, $x(t) = [x_1(t), x_2(t), \ldots, x_n(t)]^T$ is a column vector which denotes the states of neurons in the neural network dynamic system, and the elements of symmetric matrix A denote the connection weights between those neurons. We presented the schematic diagram of the neural network in Figure 3, from which we can see that it is a recurrent neural network since the input is just the output of the system.

In the practical applications, we often only need a nonzero column vector $x(0) = [x_1(0), x_2(0), \ldots, x_n(0)]^T$ to start the neural network system by the following update rule:

$$
x(j+1) = x(j) + \tau \left(x(j)^T x(j) A x(j) - x(j)^T A x(j) x(j) \right),
$$
(52)

where j denote the jth iteration and τ is a small time step. The iteration stops once $\|x(j+1) - x(j)\| < \varepsilon$, where ε is a small constraint error that can be set in advance. If $\|x(j+1) - x(j)\| < \varepsilon$, we could regard that $\|x(j+1) - x(j)\| = 0$, that is, $x(j)^T x(j) A x(j) - x(j)^T A x(j) x(j) = 0$; so we have $A x(j) = (x(j)^T A x(j)/x(j)^T x(j)) x(j)$, according to the theory in [23], $x(j)$ is the eigenvector corresponding to the modulus largest eigenvalue which can be denoted as $x(j)^T A x(j)/x(j)^T x(j)$.

4. Examples and Discussion

Three experiments are presented to verify our results. The following real normal matrix A (randomly generated) was used in those three experiments:

$$
A = \begin{pmatrix}
1.8818 & -2.1563 & -0.5979 & -0.4768 & 0.7270 & 0.4133 \\
-1.4267 & 0.5258 & 1.4946 & 0.7125 & 0.4700 & 2.5384 \\
0.4273 & 0.7784 & 2.8645 & 0.8587 & 1.3792 & -1.0153 \\
-0.1929 & 1.3273 & -1.3125 & 3.1629 & 0.2573 & 0.6329 \\
1.5966 & -0.6593 & -0.3799 & 0.2058 & 2.9845 & 0.6404 \\
1.0627 & 1.9973 & -0.0249 & -1.5727 & 0.9289 & 2.3259
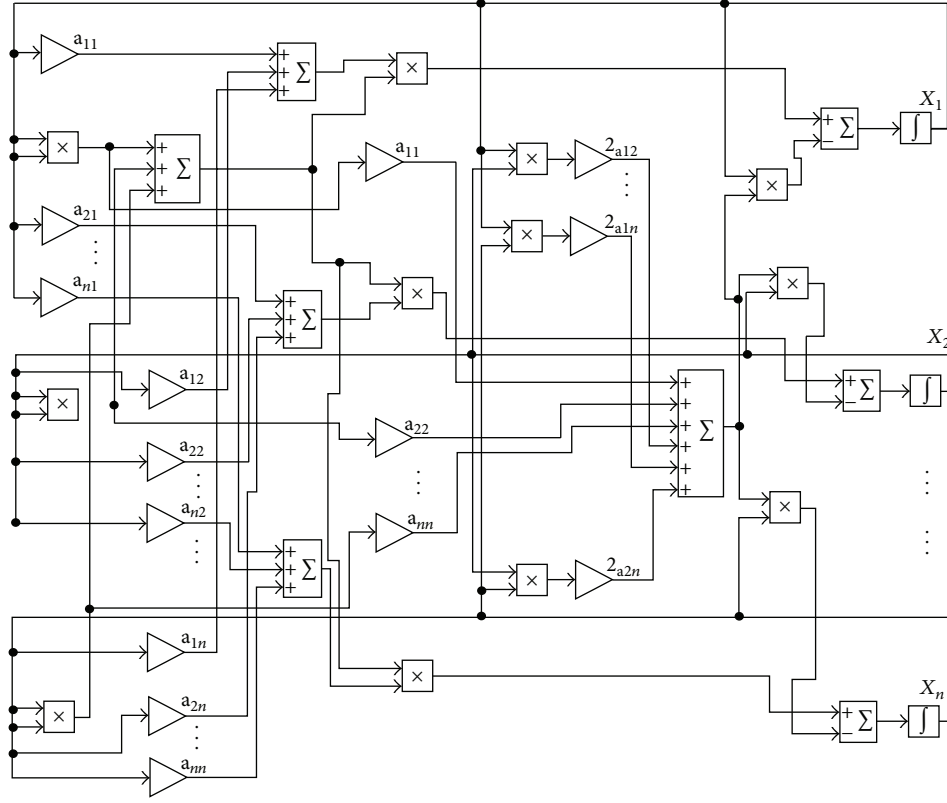\end{pmatrix}.
$$
(53)

FIGURE 3: The schematic diagram of the neural network equation (2) for solving the eigenvector corresponding to the modulus largest eigenvalues of real symmetric matrix A.

Using the $[V, D] = \mathrm{eig}(A)$ function in Matlab, we got the eigenvalues of A as $\lambda_1 = -2.3420$, $\lambda_2 = 1.005$, $\lambda_3 = \overline{\lambda_4} = 3.1967 + 1.7860i$, and $\lambda_5 = \overline{\lambda_6} = 4.3445 + 1.2532i$, as well as the corresponding eigenvectors $u_1, \ldots, u_6$ as follows:

$$u_1 = \begin{pmatrix} 0.3754 \\ 0.7510 \\ -0.2146 \\ -0.1664 \\ 0.0280 \\ -0.4696 \end{pmatrix},$$

$$u_2 = \begin{pmatrix} 0.6657 \\ -0.0242 \\ 0.2994 \\ 0.2594 \\ -0.5888 \\ 0.2295 \end{pmatrix},$$

$$u_3 = \begin{pmatrix} -0.0294 + 0.1297i \\ -0.0243 - 0.1385i \\ 0.6165 \\ -0.1917 + 0.5472i \\ 0.0913 + 0.2785i \\ -0.2708 - 0.2950i \end{pmatrix},$$

$$u_5 = \begin{pmatrix} 0.3207 + 0.2956i \\ 0.0472 - 0.4424i \\ 0.0711 - 0.2167i \\ -0.1146 - 0.3213i \\ 0.4903 \\ 0.3692 - 0.2583i \end{pmatrix},$$

$$u_4 = \overline{u}_3,$$

$$u_6 = \overline{u}_5. \tag{54}$$

We can see that all of (C1) to (C6) hold except (C2). For simplicity, denote $\lim_{t \to \infty} x(t)$ by $x(\infty)$.

Example 19 (for Section 2.1). We used (10) with the following initial condition (randomly generalized)

$$x(0) = [0.2329, -0.0023, -0.8601, \\ -0.2452, -0.1761, -0.1577]^T \tag{55}$$

to get $|b_3|$ (the largest absolute value of the imaginary part among λ_i) that $\sqrt{-x(t)^T (A - A^T)^2 x(t) / 4x(t)^T x(t)}$ should converge to. From Lemma 5, $x(t)^T (A + A^T)x(t)/2x(t)^T x(t) \to a_3$ should hold. By Lemma 7, $x(t)$ and $(A - A^T)x(t)/2|b_3|$ should converge to the imaginary
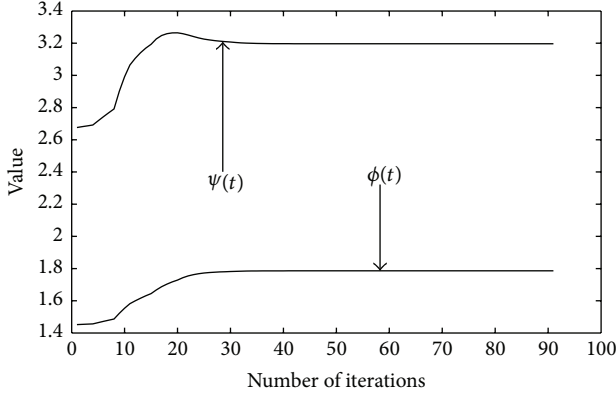
FIGURE 4: Trajectories of $\phi(t) = \sqrt{-x(t)^T(A - A^T)^2 x(t)/4x(t)^T x(t)}$ and $\psi(t) = x(t)^T(A + A^T)x(t)/2x(t)^T x(t)$ based on (10) with initial $x(0)$ as (55), which should converge to $|b_3|$ and a_3, respectively.
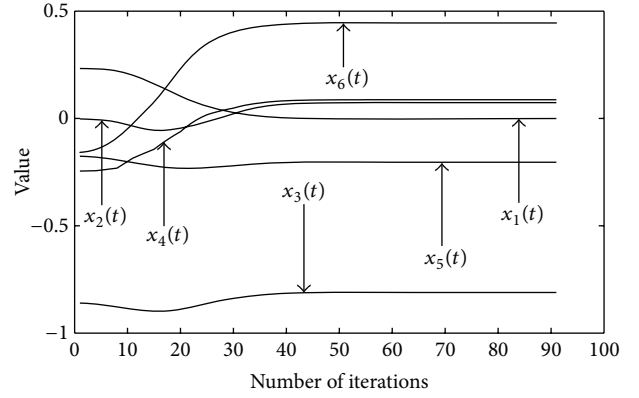

FIGURE 5: Trajectories of $x(t)$, the solution of (10) with initial $x(0)$ as (55), which should converge to the imaginary part of an eigenvector of A corresponding to the eigenvalue λ_3.

and real parts of an eigenvector corresponding to λ_3, respectively. The transient behaviors of the above four variables are shown in Figures 4, 5, and 6, respectively. After convergence, we saw

$$\frac{\left(A - A^T\right)x(\infty)}{2|b_3|} + x(\infty)i$$

$$= \begin{pmatrix} 0.1795 - 0.0005i \\ -0.1747 + 0.0737i \\ -0.1857 - 0.8107i \\ 0.7772 + 0.0872i \\ 0.3386 - 0.2039i \\ -0.3064 + 0.4450i \end{pmatrix} \tag{56}$$
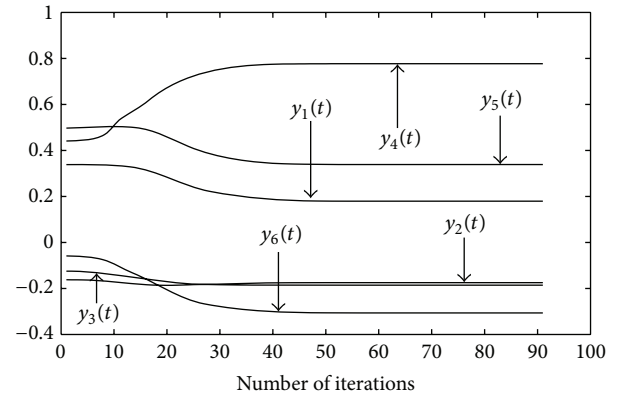
$$= (-0.3012 - 1.3149i)\, u_3.$$

FIGURE 6: Trajectories of $y(t) = (A - A^T)x(t)/2b_3$ based on (10) with initial $x(0)$ as (55), which should converge to the real part of an eigenvector of A corresponding to the eigenvalue λ_3.

Thus, the estimated complex vector is an eigenvector of A corresponding to λ_3.

Although we can use (25) to get the smallest absolute value of the imaginary part among λ_i (it is zero in this experiment), neither the corresponding real part nor the eigenvector can be obtained from Lemmas 8 or 9 since (C2) does not hold.

Example 20 (for Section 2.2). We used (33) with the same $x(0)$ as (55) to get a_5 (the largest real part among λ_i) that $x(t)^T(A + A^T)x(t)/2x(t)^T x(t)$ should converge to. Based on Lemma 12, $\sqrt{(x(t)^T(AA^T)x(t)/x(t)^T x(t)) - a_5^2} \rightarrow |b_5|$ should hold. The transient behaviors of such two variables are shown in Figure 7. Since $b_5 \neq 0$, from Lemma 13, $x(t)$ and $(A - A^T)x(t)/2|b_5|$ should converge to the imaginary and real parts of an eigenvector corresponding to λ_5, as shown

in Figures 8 and 9. After convergence, we saw

$$\frac{\left(A - A^T\right)x(\infty)}{2|b_5|} + x(\infty)i$$

$$= \begin{pmatrix} -0.5695 - 0.1474i \\ 0.2254 + 0.5561i \\ 0.0534 + 0.3029i \\ 0.3407 + 0.3091i \\ -0.5830 + 0.3120i \\ -0.2746 + 0.5420i \end{pmatrix} \tag{57}$$

$$= (-1.1890 + 0.6363i)\, u_5.$$

Hence, the estimated complex vector is an eigenvector of A corresponding to λ_5.

Based on (37), we got a_1 (the smallest real part among λ_i). After convergence, we saw that $x(\infty)^T(A + A^T)x(\infty)/2x(\infty)^T x(\infty) = a_1 = \lambda_1 = -2.3420$, $\sqrt{(x(\infty)^T(AA^T)x(\infty)/x(\infty)^T x(\infty)) - a_1^2} = |b_1| = 0$, and
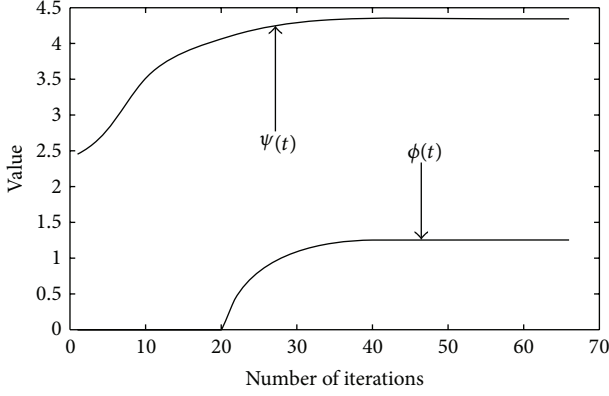
FIGURE 7: Trajectories of $\psi(t) = x(t)^T(A + A^T)x(t)/2x(t)^Tx(t)$ and $\phi(t) = \sqrt{(x(t)^T(AA^T)x(t)/x(t)^Tx(t)) - a_5^2}$ based on (33) with initial $x(0)$ as (55), which should converge to a_5 and $|b_5|$, respectively.
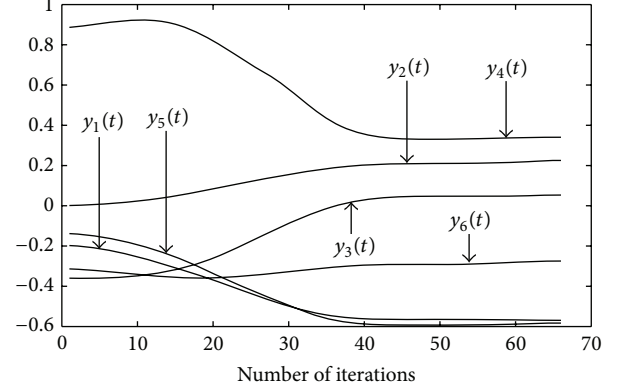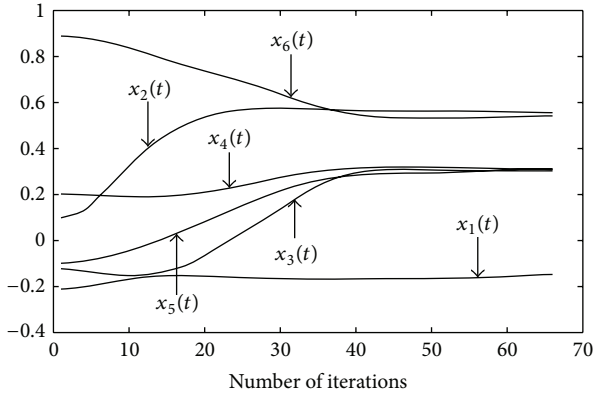

FIGURE 8: Trajectories of $x(t)$, the solution of (33) with initial $x(0)$ as (55), which should converge to the imaginary part of an eigenvector of A corresponding to the eigenvalue λ_5.

FIGURE 9: Trajectories of $y(t) = (A - A^T)x(t)/2|b_5|$ based on (33) with initial $x(0)$ as (55), which should converge to the real part of an eigenvector of A corresponding to the eigenvalue λ_5.
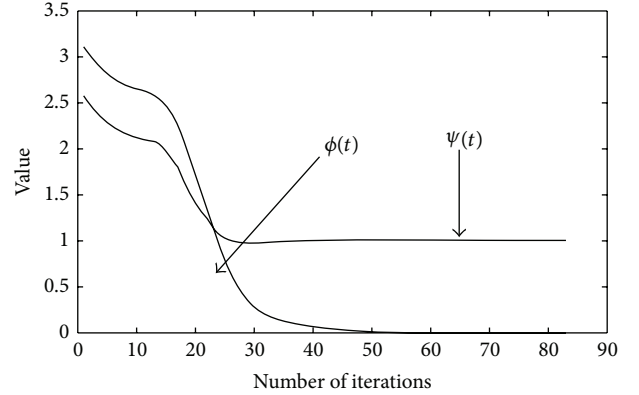

FIGURE 10: Trajectories of $\psi(t) = x(t)^T(A + A^T)x(t)/2x(t)^Tx(t)$ and $\phi(t) = \sqrt{(x(t)^T(AA^T)x(t)/x(t)^Tx(t)) - a_2^2}$ based on (40) with initial $x(0)$ as (55), which should converge to a_2 and $|b_2|$, respectively.

$x(\infty)$ was equal to

$$x(\infty) = \begin{pmatrix} -0.3581 \\ -0.7165 \\ 0.2048 \\ 0.1588 \\ -0.0267 \\ 0.4480 \end{pmatrix} = -0.9541u_1, \qquad (58)$$

just as expected from Lemma 14.

Example 21 (for Section 2.3). Based on (39) and Lemma 16, we can get λ_5 and one corresponding eigenvector again because $|\lambda_5|$ is the largest modulus of λ_i. In addition, we used (40) with the same $x(0)$ as (55) to get $|\lambda_2|$, the smallest modulus of λ_i. By Lemma 17, $a_2 = x(\infty)^T(A + A^T)x(\infty)/2x(\infty)^Tx(\infty)$ and $|b_2| = \sqrt{(x(\infty)^T(AA^T)x(\infty)/x(\infty)^Tx(\infty)) - a_2^2}$ should hold. The transient behaviors of such two variables are shown in Figure 10. After convergence, we saw that $b_2 = 0$. Hence, from Lemma 17, $x(t)$ should converge to a constant multiple

of u_2, which is shown in Figure 11. As expected, $x(\infty)$ was equal to

$$x(\infty) = \begin{pmatrix} -0.6337 \\ 0.0230 \\ -0.2850 \\ -0.2469 \\ 0.5605 \\ -0.2185 \end{pmatrix} = -0.9519u_2. \qquad (59)$$

Example 22 (for extension to arbitrary real matrices). According to the theory in Section 2.1, we present here an experiment for arbitrary real matrices to verify the effectiveness of the norm-reducing technique in Section 2.4. Considering the following nonnormal matrix A:

$$A = \begin{pmatrix} 3.4633 & -3.7487 & 2.9281 & -0.5106 & -0.0455 \\ 4.4860 & 3.4259 & -3.2010 & 5.0625 & 1.2120 \\ 0.5268 & -1.0311 & 1.5386 & -0.4766 & -1.4478 \\ -2.4768 & 1.2585 & 2.0525 & 2.3539 & -0.8487 \\ 0.0498 & -3.1267 & -6.5297 & -1.5106 & 2.2083 \end{pmatrix}, \qquad (60)$$
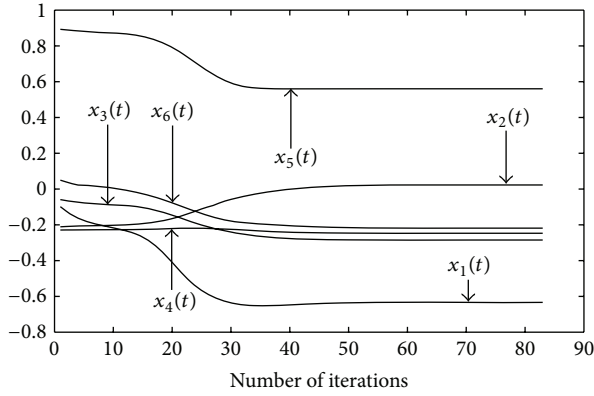
FIGURE 11: Trajectories of $x(t)$, the solution of (40), with initial $x(0)$ as (55), which should converge to an eigenvector of A corresponding to the eigenvalue λ_2.
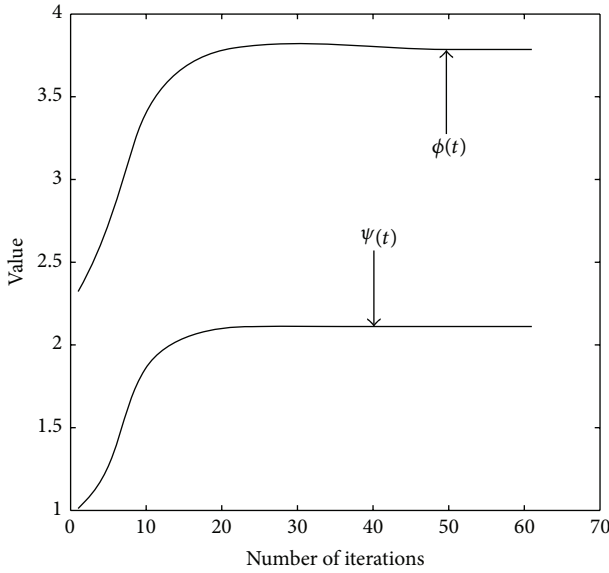

FIGURE 12: Trajectories of $\phi(t) = \sqrt{-x(t)^T(A_p - A_p^T)^2 x(t)/4x(t)^T x(t)}$ and $\psi(t) = x(t)^T(A_p + A_p^T)x(t)/2x(t)^T x(t)$ based on (10) with initial $x(0)$ as (65), which should converge to $|b_2|$ and a_2, respectively.
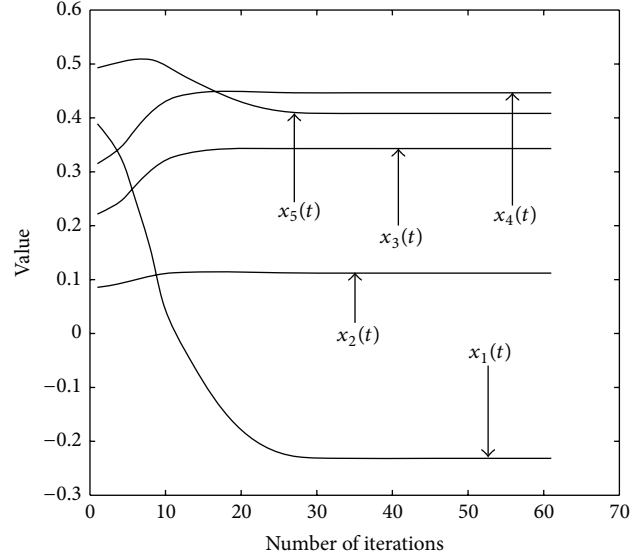

FIGURE 13: Trajectories of $x(t)$, the solution of (10) with initial $x(0)$ as (65), which should converge to the imaginary part of an eigenvector of A_p corresponding to the eigenvalue λ_2.
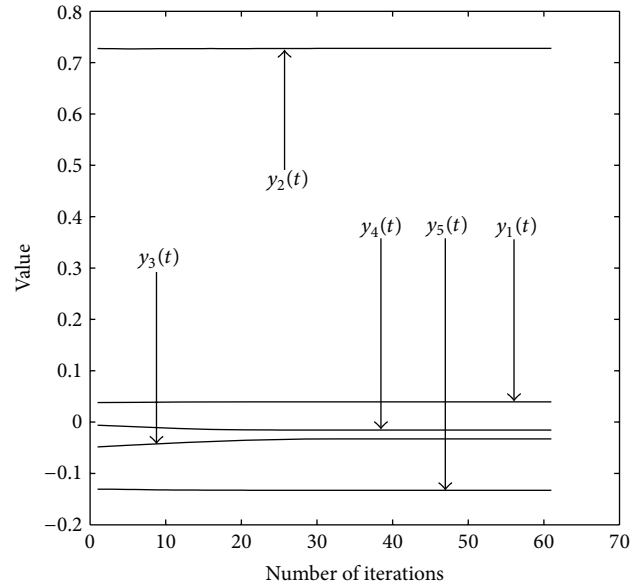

FIGURE 14: Trajectories of $y(t) = (A_p - A_p^T)x(t)/2b_2$ based on (10) with initial $x(0)$ as (65), which should converge to the real part of an eigenvector of A_p corresponding to the eigenvalue λ_2.

the Frobenius norm of A is $\|A\| = \sqrt{187.8869} > \sqrt{\sum_{i=1}^{5} \lambda_i^2} = 97.1780$, where λ_i, $i = 1, 2, \ldots, 5$ are the eigenvalues of matrix A, so matrix A is a nonnormal matrix. According to the theory in Section 2.4, the eigenpairs problem of nonnormal real matrix can be converted into the eigenpairs problem of the corresponding normal matrix A_p, which can be calculated by norm-reducing technique:

$$A_p = \begin{pmatrix} 2.3438 & 0.7915 & 1.1071 & 1.8624 & -2.6867 \\ -1.5918 & 2.1016 & 1.7249 & 2.2067 & 1.9843 \\ 0.8659 & -1.7434 & 4.5103 & -1.2708 & 0.0408 \\ 1.7777 & -2.3015 & -0.9898 & 3.5815 & 0.7201 \\ -2.4688 & -2.3136 & -0.4746 & 0.0545 & 0.4528 \end{pmatrix},$$

(61)

and the corresponding P_∞ as follows:

$$P_\infty = \begin{pmatrix} 28.2974 & -28.8442 & 115.6698 & 16.4765 & 13.5575 \\ -35.9689 & 48.5475 & 4.3465 & 58.1384 & 9.0642 \\ -10.2373 & -14.2976 & 14.1168 & 23.6416 & -40.2882 \\ -3.4360 & -25.4122 & -114.2417 & 15.0191 & 15.1720 \\ -79.1532 & -35.3465 & 75.4347 & -24.8455 & -48.2893 \end{pmatrix},$$

(62)

which satisfies the relationship $A_p = P_\infty^{-1} A P_\infty$. In order to solve the eigenpairs of matrix A, we have to calculate the eigenpairs of the matrix A_p at first. Direct calculations

of the eigenvalues of matrix A_p are $\lambda_1 = -1.8342$, $\lambda_2 = \overline{\lambda_3} = 2.1114 + 3.786i$, $\lambda_4 = \overline{\lambda_5} = 5.3007 + 0.1328i$ and those are just the eigenvalues of nonnormal matrix A and the corresponding eigenvectors as follows:

$$
u_1 = \begin{pmatrix} 0.6079 \\ 0.0886 \\ -0.1195 \\ -0.2805 \\ 0.7278 \end{pmatrix}, \quad u_2 = \begin{pmatrix} 0.0035 - 0.2240i \\ 0.7020 \\ 0.0189 + 0.3280i \\ 0.0502 + 0.4231i \\ -0.0660 + 0.4040i \end{pmatrix} = \overline{u_3},
$$

$$
u_4 = \begin{pmatrix} 0.1769 - 0.4835i \\ -0.0129 + 0.0561i \\ 0.6204 \\ -0.2522 - 0.4643i \\ -0.1415 + 0.2181i \end{pmatrix} = \overline{u_5}.
$$

(63)

The eigenvectors of nonnormal matrix A, which corresponding to the eigenvalues with largest imaginary parts in absolute are

$$
u_2^A = \begin{pmatrix} -0.2353 - 0.5331i \\ -0.6416 \\ 0.0449 - 0.0581i \\ 0.3856 + 0.0270i \\ 0.0747 - 0.2973i \end{pmatrix} = \overline{u_3^A}.
$$

(64)

According to the results above, the eigenvalues with largest imaginary in absolute are λ_2 and λ_3, and let $a_2 \pm |b_2|i$ denote them.

We used (10) with the following initial condition (randomly generalized):

$$
x(0) = [0.3891, 0.0856, 0.2215, 0.3151, 0.4929]^T
$$

(65)

to get $|b_2|$ (the largest absolute value of the imaginary part among λ_i) that $\sqrt{-x(t)^T (A_p - A_p^T)^2 x(t)/4x(t)^T x(t)}$ should converge to. From Lemma 5, $x(t)^T (A_p + A_p^T)x(t)/2x(t)^T x(t) \rightarrow a_2$ should hold. By Lemma 7, $x(t)$ and $(A_p - A_p^T)x(t)/2|b_2|$ should converge to the imaginary and real parts of an eigenvector corresponding to λ_2, respectively. The transient behaviors of the above four variables are shown in Figures 12, 13, and 14, respectively.

After convergence, we saw

$$
\frac{\left(A_p - A_p^T\right)x(\infty)}{2|b_2|} + x(\infty)i
$$

$$
= \begin{pmatrix} 0.0394 - 0.2317i \\ 0.7277 + 0.1121i \\ -0.0327 + 0.3430i \\ -0.0155 + 0.4466i \\ -0.1329 + 0.4083i \end{pmatrix}
$$

(66)

$$
= (1.0367 + 0.1597i)\, u_2.
$$

Thus, the estimated complex vector is an eigenvector of A_p corresponding to λ_2. According to the theory in Section 2.4,

the estimated complex eigenvector of nonnormal matrix A corresponding to the eigenvalue λ_2 should be

$$
P_\infty \left(\frac{\left(A_p - A_p^T\right)x(\infty)}{2|b_2|} + x(\infty)i \right)
$$

$$
= \begin{pmatrix} -25.7189 + 42.7862i \\ 31.6606 + 44.9313i \\ -6.2833 - 0.2807i \\ -17.1372 - 28.3402i \\ -24.5113 + 9.4386i \end{pmatrix}
$$

(67)

$$
= (-49.3476 - 70.0318i)\, u_2^A,
$$

from which we can see that the estimated complex vector is an eigenvector of A corresponding to λ_2.

5. Conclusion

This paper introduces a neural network based approach for computing eigenvectors of real normal matrices and the corresponding eigenvalues that have the largest or smallest modulus, have the largest or smallest real part, and have the largest or smallest imaginary part in absolute value. All the computation can be carried out in real vector space although eigenpairs may be complex, which can reduce the scale of networks a lot. We also shed light on extending this method to the case of general real matrices by employing the norm-reducing technique proposed in other literatures. Four simulation examples verified the validity of our proposed algorithm.

References

[1] S. Attallah and K. Abed-Meraim, "A fast adaptive algorithm for the generalized symmetric eigenvalue problem," *IEEE Signal Processing Letters*, vol. 15, pp. 797–800, 2008.

[2] T. Laudadio, N. Mastronardi, and M. Van Barel, "Computing a lower bound of the smallest eigenvalue of a symmetric positive-definite Toeplitz matrix," *IEEE Transactions on Information Theory*, vol. 54, no. 10, pp. 4726–4731, 2008.

[3] J. Shawe-Taylor, C. K. I. Williams, N. Cristianini, and J. Kandola, "On the eigenspectrum of the Gram matrix and the generalization error of kernel-PCA," *IEEE Transactions on Information Theory*, vol. 51, no. 7, pp. 2510–2522, 2005.

[4] D. Q. Wang and M. Zhang, "A new approach to multiple class pattern classification with random matrices," *Journal of Applied Mathematics and Decision Sciences*, no. 3, pp. 165–175, 2005.

[5] M. R. Bastian, J. H. Gunther, and T. K. Moon, "A simplified natural gradient learning algorithm," *Advances in Artificial Neural Systems*, vol. 2011, Article ID 407497, 9 pages, 2011.

[6] T. H. Le, "Applying artificial neural networks for face recognition," *Advances in Artificial Neural Systems*, vol. 2011, Article ID 673016, 16 pages, 2011.

[7] Y. Xia, "An extended projection neural network for constrained optimization," *Neural Computation*, vol. 16, no. 4, pp. 863–883, 2004.

[8] Y. Xia and J. Wang, "A recurrent neural network for solving nonlinear convex programs subject to linear constraints," *IEEE*

Transactions on Neural Networks, vol. 16, no. 2, pp. 379–386, 2005.

[9] T. Voegtlin, "Recursive principal components analysis," *Neural Networks*, vol. 18, no. 8, pp. 1051–1063, 2005.

[10] J. Qiu, H. Wang, J. Lu, B. Zhang, and K.-L. Du, "Neural network implementations for PCA and its extensions," *ISRN Artificial Intelligence*, vol. 2012, Article ID 847305, 19 pages, 2012.

[11] H. Liu and J. Wang, "Integrating independent component analysis and principal component analysis with neural network to predict Chinese stock market," *Mathematical Problems in Engineering*, vol. 2011, Article ID 382659, 15 pages, 2011.

[12] F. L. Luo, R. Unbehauen, and A. Cichocki, "A minor component analysis algorithm," *Neural Networks*, vol. 10, no. 2, pp. 291–297, 1997.

[13] G. H. Golub and C. F. Van Loan, *Matrix Computations*, vol. 3 of *Johns Hopkins Series in the Mathematical Sciences*, Johns Hopkins University Press, Baltimore, Md, USA, 1983.

[14] C. Chatterjee, V. P. Roychowdhury, J. Ramos, and M. D. Zoltowski, "Self-organizing algorithms for generalized eigen-decomposition," *IEEE Transactions on Neural Networks*, vol. 8, no. 6, pp. 1518–1530, 1997.

[15] H. Kakeya and T. Kindo, "Eigenspace separation of auto-correlation memory matrices for capacity expansion," *Neural Networks*, vol. 10, no. 5, pp. 833–843, 1997.

[16] Y. Liu, Z. You, and L. Cao, "A functional neural network for computing the largest modulus eigenvalues and their corresponding eigenvectors of an anti-symmetric matrix," *Neurocomputing*, vol. 67, no. 1–4, pp. 384–397, 2005.

[17] Y. Liu, Z. You, and L. Cao, "A functional neural network computing some eigenvalues and eigenvectors of a special real matrix," *Neural Networks*, vol. 18, no. 10, pp. 1293–1300, 2005.

[18] Y. Liu, Z. You, and L. Cao, "A recurrent neural network computing the largest imaginary or real part of eigenvalues of real matrices," *Computers & Mathematics with Applications*, vol. 53, no. 1, pp. 41–53, 2007.

[19] E. Oja, "Principal components, minor components, and linear neural networks," *Neural Networks*, vol. 5, no. 6, pp. 927–935, 1992.

[20] R. Perfetti and E. Massarelli, "Training spatially homogeneous fully recurrent neural networks in eigenvalue space," *Neural Networks*, vol. 10, no. 1, pp. 125–137, 1997.

[21] L. Xu, E. Oja, and C. Y. Suen, "Modified Hebbian learning for curve and surface fitting," *Neural Networks*, vol. 5, no. 3, pp. 441–457, 1992.

[22] Q. Zhang and Y. W. Leung, "A class of learning algorithms for principal component analysis and minor component analysis," *IEEE Transactions on Neural Networks*, vol. 11, no. 2, pp. 529–533, 2000.

[23] Y. Zhang, Y. Fu, and H. J. Tang, "Neural networks based approach for computing eigenvectors and eigenvalues of symmetric matrix," *Computers & Mathematics with Applications*, vol. 47, no. 8-9, pp. 1155–1164, 2004.

[24] L. Mirsky, "On the minimization of matrix norms," *The American Mathematical Monthly*, vol. 65, pp. 106–107, 1958.

[25] C. P.] Huang and R. T. Gregory, *A Norm-Reducing Jacobi-Like Algorithm for the Eigenvalues of Non-Normal Matrices*, Colloquia Mathematica Societatis Janos Bolyai, Keszthely, Hungary, 1977.

[26] I. Kiessling and A. Paulik, "A norm-reducing Jacobi-like algorithm for the eigenvalues of non-normal matrices," *Journal of Computational and Applied Mathematics*, vol. 8, no. 3, pp. 203–207, 1982.

[27] R. Sacks-Davis, "A real norm-reducing Jacobi-type eigenvalue algorithm," *The Australian Computer Journal*, vol. 7, no. 2, pp. 65–69, 1975.