Computers and Mathematics with Applications 60 (2010) 1385–1392

Contents lists available at ScienceDirect

Computers and Mathematics with Applications

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

Another neural network based approach for computing eigenvalues and eigenvectors of real skew-symmetric matrices

Ying Tang *, Jianping Li

School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu 610054, China

ARTICLE INFO

Article history:
Received 25 May 2009
Received in revised form 15 June 2010
Accepted 15 June 2010

Keywords:
Neural network
Real skew-symmetric matrix
Eigenvalue
Eigenvector

ABSTRACT

This paper introduces a novel neural network based approach for extracting the eigenvalues with the largest or smallest modulus of real skew-symmetric matrices, as well as the corresponding eigenvectors. To this end, unlike the previous neural network based methods that can be summarized by some $2n$-dimensional ordinary differential equations (ODEs), where n is the order of the given skew-symmetric matrix, our proposed approach corresponds to an ODE of order n, instead of $2n$. Hence, the scale of networks can be reduced a lot. Simulations verify the computational capability of such approach.

1. Introduction

Computation of eigenvalues and the corresponding eigenvectors has been an attractive topic for a long time, which is important both in theory and in many engineering fields such as image compression and signal processing, etc. Lots of neural network based methods have been proposed for solving this problem [1–16]. Two excellent review articles of this topic can be found in [17,18].

However, most of those studies focused on computing eigenvalues and the corresponding eigenvectors of real symmetric matrices. The following two 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 well-studied in [3] and [9], respectively, where $x(t) \in R^n$ and A is a real symmetric matrix. Both (1) and (2) are efficient for computing the largest eigenvalue and the corresponding eigenvector. Moreover, they can succeed in computing the smallest eigenvalue and the corresponding eigenvector by simply replacing A with $-A$.

Then, some extensions as regards this topic appeared. The following ODE was proposed in [8] for solving the generalized eigenvalue problem:

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = Ax(t) - f(x(t))Bx(t), \tag{3}$$

where both A and B are real symmetric matrices and f can take a general form to some degree. Note that when B is the identity matrix, (3) can also be used to solve a conventional eigenvalue problem like (1) and (2).

* Corresponding author. Tel.: +86 13558710280.
E-mail addresses: mathtygo@yahoo.com (Y. Tang), jpli2222@uestc.edu.cn (J. Li).

0898-1221/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.camwa.2010.06.020

Recently, such neural network based methods were extended to the case of special real matrices in [12,13]. In addition, two similar neural networks, which can be summarized by $2n$-dimensional ODEs, were proposed in [14,15] for computing eigenpairs of n-by-n real skew-symmetric matrices. Since the dimensionality of those two ODEs introduced in [14,15] is $2n$, much more than n, the scale of networks would be enlarged, making networks more complex. Hence, they are possibly unacceptable in practice. Due to the close relationship between skew-symmetric matrices and infinitesimal rotations [19] that are involved in many engineering applications [20,21], a simpler neural network for dealing with the case of skew-symmetric matrices is required.

In this paper, the problem of extracting eigenpairs of n-dimensional real skew-symmetric matrices is translated into that for the related real symmetric matrices by employing Householder reflector and similarity transformations. The translated problem can then be solved using such n-dimensional ODEs as (1) and (2) or (3). In the next section six lemmas are presented. The first four address the transformations involved in this paper and the last two are the required stability theorems established in [9]. Experimental results are provided in Section 3. Section 4 concludes the paper.

2. Main results

It is well-known that a Householder reflector is defined by

$$H = I_n - 2uu^T, \quad u \in R^n, \quad \|u\| = 1, \tag{4}$$

where I_n is the identity matrix of order n and $\|\cdot\|$ denotes the Frobenius norm. Clearly, H is orthogonal and symmetric. From a computational viewpoint, such transformations originate in the annihilation of selected elements of vectors or matrices and are represented by the isometric mapping of a driving vector z into a stretching of a vector of the canonical basis e_i^n (the ith column of I_n).

Lemma 1 (*cf. [22]*)**.** *Given any $z \in R^n$ and e_i^n, $i = 1, \ldots, n$, $Hz = \|z\|e_i^n$ holds if $H = I_n - 2uu^T$ and $u = \frac{z-\|z\|e_i^n}{\|z-\|z\|e_i^n\|}$.*

Lemma 2. *Given any real skew-symmetric matrix A, there exist $n-2$ orthogonal transformations P_i, $i = 1, \ldots, n-2$, such that*

$$P_{n-2}\cdots P_1 A P_1^T \cdots P_{n-2}^T = \begin{pmatrix} 0 & a_1 & & & & \\ -a_1 & 0 & a_2 & & & \\ & -a_2 & 0 & a_3 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -a_{n-2} & 0 & a_{n-1} \\ & & & & -a_{n-1} & 0 \end{pmatrix} = A_t. \tag{5}$$

Proof. Assume

$$A = \begin{pmatrix} 0 & w_1^T \\ -w_1 & A_1 \end{pmatrix}. \tag{6}$$

Here, $w_1 \in R^{n-1}$ and A_1 is a real skew-symmetric matrix of order $n-1$. By Lemma 1, there exists H_1 such that $H_1 w_1 = \|w_1\|e_1^{n-1}$. Let

$$P_1 = \begin{pmatrix} 1 & 0 \\ 0 & H_1 \end{pmatrix}. \tag{7}$$

Therefore,

$$P_1 A P_1^T = \begin{pmatrix} 0 & w_1^T H_1^T \\ -H_1 w_1 & H_1 A_1 H_1^T \end{pmatrix} = \begin{pmatrix} 0 & \|w_1\|(e_1^{n-1})^T \\ -\|w_1\|e_1^{n-1} & B_2 \end{pmatrix}, \quad B_2 = H_1 A_1 H_1^T. \tag{8}$$

Clearly, B_2 is a real skew-symmetric matrix. For $i = 2, \ldots, n-2$, let

$$B_i = \begin{pmatrix} 0 & w_i^T \\ -w_i & A_i \end{pmatrix}. \tag{9}$$

Here, $w_i \in R^{n-i}$ and A_i is a real skew-symmetric matrix of order $n-i$. On the basis of Lemma 1, choose H_i such that $H_i w_i = \|w_i\|e_1^{n-i}$ and let

$$P_i = \begin{pmatrix} I_i & 0 \\ 0 & H_i \end{pmatrix}. \tag{10}$$

Then, it is straightforward to verify that the matrix $P_{n-2}\cdots P_1 A P_1^T \cdots P_{n-2}^T$ takes the form of A_t in (5), thus proving the lemma. □

Let

$$P = P_{n-2} \cdots P_1. \tag{11}$$

Clearly, P is an orthogonal matrix. Then, the following simple lemma is given whose proof may be found in many textbooks on matrices such as [23]:

Lemma 3. *If S and W are similar matrices satisfying $T^{-1}ST = W$, the eigenvalues of S are all the same as those of W. If υ is an eigenvector of W corresponding to the eigenvalue λ, $T\upsilon$ is an eigenvector of S corresponding to the eigenvalue λ.*

Define the following real symmetric matrix:

$$A_s = \begin{pmatrix} 0 & a_1 & & & & \\ a_1 & 0 & a_2 & & & \\ & a_2 & 0 & a_3 & & \\ & & \ddots & \ddots & \ddots & \\ & & & a_{n-2} & 0 & a_{n-1} \\ & & & & a_{n-1} & 0 \end{pmatrix}, \tag{12}$$

where the a_j's, $j = 1, \ldots, n-1$, are the same as those in A_t defined as (5). Let $\mathrm{i} = \sqrt{-1}$ be the imaginary unit. Then, we have the following lemma:

Lemma 4. *A_t and $\mathrm{i}A_s$ are similar matrices.*

Proof. Let $Q = \mathbf{diag}[1, \mathrm{i}, \mathrm{i}^2, \ldots, \mathrm{i}^{n-1}]$. Then, it is straightforward to verify

$$Q^{-1}A_tQ = \mathrm{i}A_s, \tag{13}$$

thus proving the lemma. □

Next, note the following four facts. (1) Eigenvalues of the real skew-symmetric matrix A are zero or pure imaginary numbers. (2) If $\mathrm{i}\lambda$, $\lambda \in R$, is an eigenvalue of A corresponding to the eigenvector u, $-\mathrm{i}\lambda$ is another eigenvalue of A corresponding to the eigenvector $\bar{u}$, where $\bar{u}$ is the conjugate of u. (3) All the eigenvalues of the real symmetric matrix A_s are real numbers. (4) By Lemmas 2–4, the eigenvalues of $\mathrm{i}A_s$ are the same as those of A. Hence, if λ, $\lambda \in R$, is an eigenvalue of A_s, so is $-\lambda$. Therefore, we can denote the n eigenvalues of A_s and the corresponding eigenvectors as λ_j, υ_j, $j = 1, \ldots, n$, where $\lambda_j \in R$, $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|$, $\lambda_{2k-1} = -\lambda_{2k}$ and $\lambda_{2k-1} \geq 0$, $k = 1, \ldots, \lfloor \frac{n}{2} \rfloor$ ($\lfloor \frac{n}{2} \rfloor$ is the maximal integer that is no more than $\frac{n}{2}$).

Then, on the basis of Lemmas 2–4 again, it is important to observe the following key relation between the eigenpairs of A_s and A: If λ_j is an eigenvalue of A_s corresponding to the eigenvector υ_j, we know that $\pm\mathrm{i}\lambda_j$ are two eigenvalues of A corresponding to the eigenvectors $P^TQ\upsilon_j$ and $P^TQ\bar{\upsilon}_j$, respectively, where P and Q are defined in (11) and Lemma 4.

As we have introduced, many neural network based methods such as (2) have been proposed for computing the largest or smallest eigenvalues and the corresponding eigenvectors of any real symmetric matrix. Therefore, we replace A with A_s in (2), which reads

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = x(t)^Tx(t)A_sx(t) - x(t)^TA_sx(t)x(t). \tag{14}$$

The following result on the convergence of (14) can be found in [9].

Lemma 5 (*Theorem 4 in [9]*)**.** *Assume $x(0)$ is a nonzero vector in R^n which is not orthogonal to the eigensubspace corresponding to the largest eigenvalue of A_s. Then, the solution of (14) starting from $x(0)$ converges to an eigenvector corresponding to the largest eigenvalue of A_s that is equal to* $\lim_{t\to+\infty} \frac{x(t)^TA_sx(t)}{x(t)^Tx(t)}$.

Hence, using (14), we can get the eigenvectors $P^TQ\upsilon_1$ and $P^TQ\bar{\upsilon}_1$ of A and the corresponding eigenvalues $\pm\mathrm{i}\lambda_1$ that have the largest modulus.

However, if we directly replace $-A_s$ with A_s in (14), we can only get the eigenvalue λ_2 ($\lambda_2 = -\lambda_1$), which is the smallest eigenvalue of A_s and has been obtained from (14). Note that A_s^2 is also a real symmetric matrix, the eigenvalues of which are λ_j^2 corresponding to the eigenvector υ_j. Hence, to get the eigenvalues $\pm\mathrm{i}\lambda_n$ that have the smallest modulus of A and the corresponding eigenvectors $P^TQ\upsilon_n$ and $P^TQ\bar{\upsilon}_n$, we should replace A with $-A_s^2$ in (2) as follows:

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = -x(t)^Tx(t)A_s^2x(t) + x(t)^TA_s^2x(t)x(t). \tag{15}$$

Lemma 6 (*Theorem 5 in [9]*). *Assume $x(0)$ is a nonzero vector in R^n which is not orthogonal to the eigensubspace corresponding to the smallest eigenvalue of A_s^2. Then, the solution of (15) starting from $x(0)$ converges to an eigenvector corresponding to the smallest eigenvalue of A_s^2 that is equal to $\lim_{t\to+\infty}\frac{x(t)^T A_s^2 x(t)}{x(t)^T x(t)}$.*

Hence, using (15), we can get λ_n^2 and υ_n. By the above analysis, we know that $P^T Q \upsilon_n$ and $P^T Q \overline{\upsilon}_n$ are two eigenvectors of A corresponding to the two eigenvalues $\pm\mathrm{i}\lambda_n$ that have the smallest modulus, respectively.

However, only the discrete-time version of (14) and (15) can be used in practice. Then, we simply discuss some issues related to the discrete-time version of (2) such as the selection of the step length. Firstly, it is straightforward to show that the solution of (2) has invariant Frobenius norm [9]. Hence, when choosing the normalized initial condition for (2), i.e., $\|x(0)\| = 1$, there exists $x(t)^T x(t) = 1$, indicating that in this case (2) is reduced to (1). In [2], it has been proven that the discrete-time version of (1) converges if the variable step length $\eta(k)$, $k = 1, 2, \ldots$, is chosen such that its sum is divergent but its sum of squares is convergent, i.e., $\sum_{k=1}^{\infty}\eta(k) = \infty$ and $\sum_{k=1}^{\infty}\eta(k)^2 < \infty$. Therefore, such a strategy for the step length can be applied to the discrete-time version of (14) and (15). Additional information about the accuracy of neural network based eigenpair estimators for real symmetric matrices can be found in [24]. Nevertheless, many mature software systems are available for solving the convergent ODE systems, such as the $ode15s(\cdot)$ solver in Matlab that was also used in the simulation of this paper. The experimental results in the next section showed that $ode15s(\cdot)$ with default parameters performed well for computing solutions of (14) and (15).

Finally, it is easy to see that Lemmas 5 and 6 can almost hold by random selection for $x(0)$ because the projection of random $x(0)$ on the eigensubspace corresponding to the largest or smallest eigenvalues is sure to be nonzero with high probability since the dimensionality of such eigensubspace is less than that of R^n in general. Moreover, we point out that any other stochastic approximation approaches for extracting eigenpairs of real symmetric matrices can take the place of (14) and (15). Related reading on those methods can be found in [23].

3. Simulations

Two simulations are presented to verify our results. The following real skew-symmetric matrix A was used in both experiments:

$$A = \begin{pmatrix} 0 & 0.6837 & 0.0194 & 0.1856 & 0.0885 & 0.7982 & -0.4444 \\ -0.6837 & 0 & 0.3056 & 0.0305 & -0.1352 & 0.0307 & 0.0248 \\ -0.0194 & -0.3056 & 0 & 0.2671 & 0.3347 & 0.6079 & 0.5217 \\ -0.1856 & -0.0305 & -0.2671 & 0 & -0.5776 & 0.4037 & 0.7374 \\ -0.0885 & 0.1352 & -0.3347 & 0.5776 & 0 & -0.2048 & -0.4198 \\ -0.7982 & -0.0307 & -0.6079 & -0.4037 & 0.2048 & 0 & 0.4930 \\ 0.4444 & -0.0248 & -0.5217 & -0.7374 & 0.4198 & -0.4930 & 0 \end{pmatrix}.$$

On the basis of the approach presented in Lemma 2, the orthogonal matrix P can be obtained as follows:

$$P = \begin{pmatrix} 1.0000 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.5896 & 0.0167 & 0.1601 & 0.0763 & 0.6883 & -0.3832 \\ 0 & -0.0199 & -0.1875 & 0.1253 & -0.3278 & 0.4680 & 0.7888 \\ 0 & 0.2573 & -0.5106 & -0.7618 & 0.2800 & 0.0048 & 0.1197 \\ 0 & 0.6195 & -0.3365 & 0.3073 & -0.3592 & -0.5263 & 0.0498 \\ 0 & 0.2177 & 0.5724 & -0.5278 & -0.5882 & -0.0174 & -0.0088 \\ 0 & 0.3932 & 0.5128 & 0.0729 & 0.5773 & -0.1727 & 0.4626 \end{pmatrix}.$$

Then, we can compute

$$A_s = Q^{-1}PAP^TQ = \begin{pmatrix} 0 & 1.1596 & 0 & 0 & 0 & 0 & 0 \\ 1.1596 & 0 & 0.5688 & 0 & 0 & 0 & 0 \\ 0 & 0.5688 & 0 & 1.2071 & 0 & 0 & 0 \\ 0 & 0 & 1.2071 & 0 & 0.3067 & 0 & 0 \\ 0 & 0 & 0 & 0.3067 & 0 & 0.6373 & 0 \\ 0 & 0 & 0 & 0 & 0.6373 & 0 & 0.3999 \\ 0 & 0 & 0 & 0 & 0 & 0.3999 & 0 \end{pmatrix}.$$

Using the $[V, D] = eig(\cdot)$ function in Matlab, we got the true eigenvalues of A as $\pm 1.5195\mathrm{i}$, $\pm 0.9850\mathrm{i}$, $\pm 0.7115\mathrm{i}$, 0, (i.e., $\lambda_1 = -\lambda_2 = 1.5195$, $\lambda_3 = -\lambda_4 = 0.9850$, $\lambda_5 = -\lambda_6 = 0.7115$, $\lambda_7 = 0$) and the corresponding eigenvectors $u_1, \ldots, u_7$, where $u_2 = \overline{u}_1$, $u_4 = \overline{u}_3$, $u_6 = \overline{u}_5$,

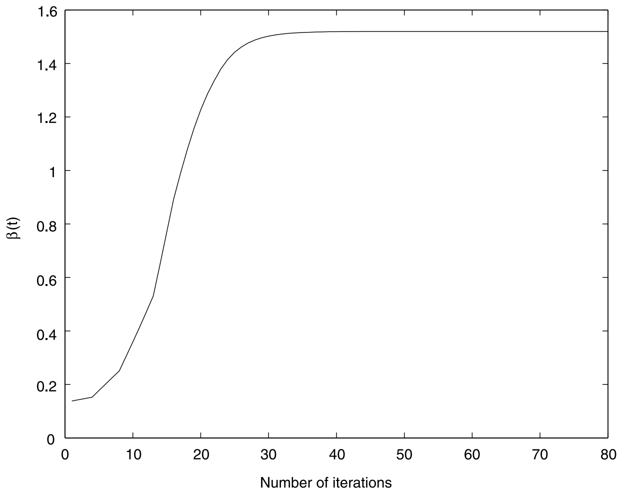

Fig. 1. Transient behavior of $\beta(t)$ based on (14) with initial $x(0)$ as in (16), which should converge to λ_1, i.e., the largest modulus of the eigenvalues of A.

$$u_1 = \begin{pmatrix} 0.3428 - 0.1929\mathrm{i} \\ 0.1638 + 0.1280\mathrm{i} \\ 0.1923 + 0.2189\mathrm{i} \\ 0.1746 + 0.3842\mathrm{i} \\ 0.0574 - 0.1781\mathrm{i} \\ -0.1149 + 0.4681\mathrm{i} \\ -0.5212 \end{pmatrix}, \qquad u_3 = \begin{pmatrix} 0.5519 \\ -0.1537 + 0.3488\mathrm{i} \\ 0.0894 - 0.3674\mathrm{i} \\ -0.0453 - 0.1664\mathrm{i} \\ 0.0582 + 0.3048\mathrm{i} \\ 0.2870 + 0.3290\mathrm{i} \\ 0.2756 - 0.1206\mathrm{i} \end{pmatrix},$$

$$u_5 = \begin{pmatrix} 0.1625 - 0.1101\mathrm{i} \\ 0.2856 + 0.1308\mathrm{i} \\ -0.1828 + 0.4613\mathrm{i} \\ -0.1407 - 0.4937\mathrm{i} \\ -0.5807 \\ -0.0311 + 0.1084\mathrm{i} \\ 0.0250 - 0.0503\mathrm{i} \end{pmatrix}, \qquad u_7 = \begin{pmatrix} 0.0655 \\ 0.6523 \\ 0.2277 \\ 0.2393 \\ 0.2511 \\ -0.3588 \\ 0.5190 \end{pmatrix}.$$

Example 1.

We use (14) with random initial

$$x(0) = [-0.0462, -0.0179, -0.1190, -0.3571, 0.8038, 0.4881, -0.6095]^T, \tag{16}$$

to find the two eigenvalues $\pm \mathrm{i}\lambda_1$ of A and the corresponding eigenvectors $P^T Q v_1$ and $P^T Q \overline{v}_1$. For simplicity, let

$$\beta(t) = \frac{x(t)^T A_s x(t)}{x(t)^T x(t)}. \tag{17}$$

By the previous analysis, we know that $\lim_{t\to+\infty} \beta(t) = \lambda_1$ and $\lim_{t\to+\infty} P^T Qx(t) = P^T Q v_1$, i.e., an eigenvector of A corresponding to the eigenvalue $\mathrm{i}\lambda_1$. The transient behavior of $\beta(t)$ and that of the real part and the imaginary part of $P^T Qx(t)$ are shown in Figs. 1, 2 and 3, respectively.

After convergence, we can see that $\beta(t) \to \lambda_1$ and the estimated complex vector is just a constant multiple of u_1 as follows:

$$\text{the estimated for } u_1 = \begin{pmatrix} -0.4655 \\ -0.0947 - 0.2271\mathrm{i} \\ -0.0713 - 0.3373\mathrm{i} \\ 0.0429 - 0.4975\mathrm{i} \\ -0.1625 + 0.1504\mathrm{i} \\ 0.3901 - 0.4160\mathrm{i} \\ 0.5375 + 0.3024\mathrm{i} \end{pmatrix} = (-1.0312 - 0.5802\mathrm{i}) \cdot u_1,$$

meaning that the estimated vector is an eigenvector of A corresponding to the eigenvalue $\mathrm{i}\lambda_1$.

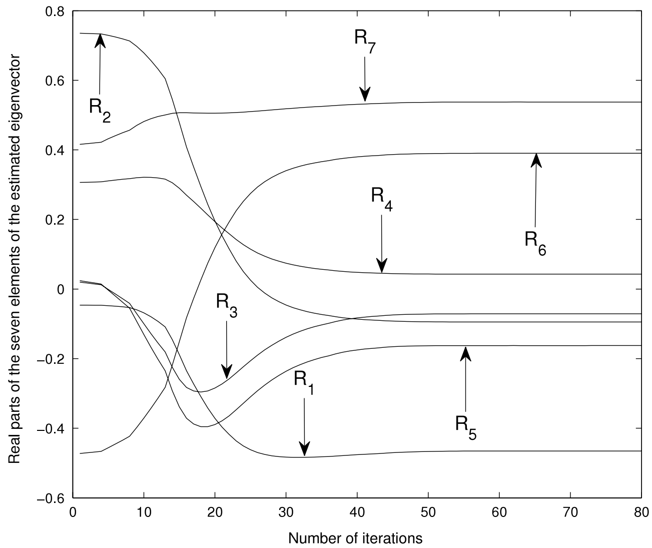

Fig. 2. Transient behavior of the real part of $P^TQx(t)$, denoted as $R(t)$, where $x(t)$ is the solution of (14) with initial $x(0)$ as in (16). $R(t)$ should converge to the real part of an eigenvector of A corresponding to the eigenvalue $\mathrm{i}\lambda_1$.

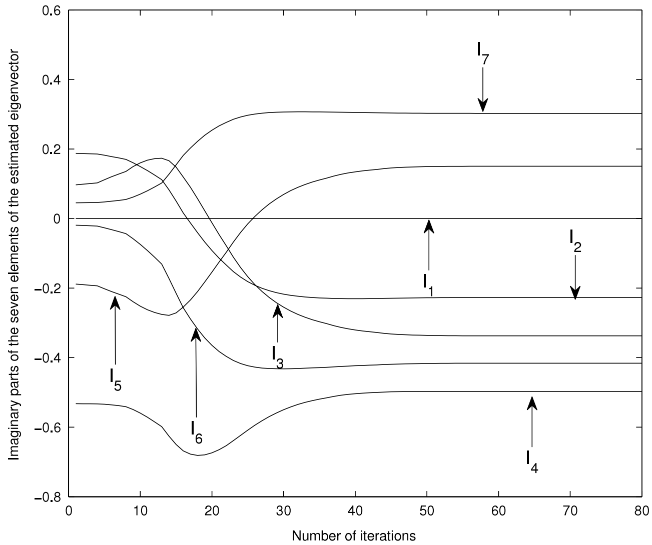

Fig. 3. Transient behavior of the imaginary part of $P^TQx(t)$, denoted as $I(t)$, where $x(t)$ is the solution of (14) with initial $x(0)$ as in (16). $I(t)$ should converge to the imaginary part of an eigenvector of A corresponding to the eigenvalue $\mathrm{i}\lambda_1$.

Example 2.

We use (15) with the same initial $x(0)$ as in (16) to compute λ_7 (the smallest modulus eigenvalue of A) and the corresponding eigenvectors P^TQv_7. For simplicity, let

$$\gamma(t) = \sqrt{\frac{x(t)^TA_s^2x(t)}{x(t)^Tx(t)}}. \tag{18}$$

By the previous analysis, we know that $\lim_{t\to+\infty}\gamma(t) = \lambda_7$ and $\lim_{t\to+\infty}P^TQx(t) = P^TQv_7$, i.e., an eigenvector of A corresponding to the eigenvalue $\mathrm{i}\lambda_7$. The transient behavior of $\gamma(t)$ and that of the real part and the imaginary part of $P^TQx(t)$ are shown in Figs. 4, 5 and 6, respectively.

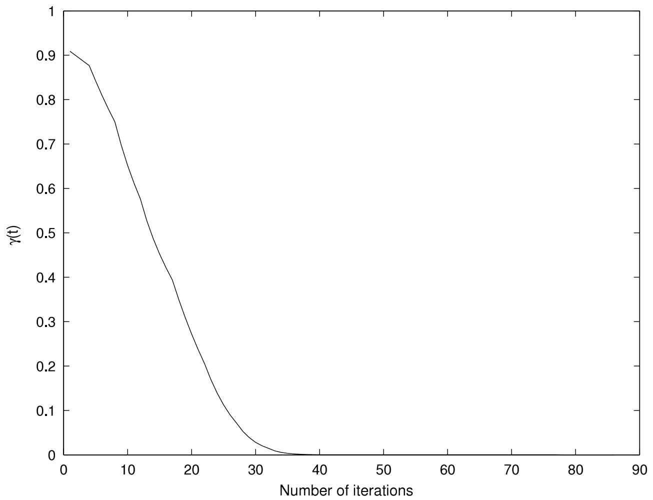

Fig. 4. Transient behavior of $\gamma(t)$ based on (15) with initial $x(0)$ as in (16), which should converge to λ_7, i.e., the smallest modulus of the eigenvalues of A.

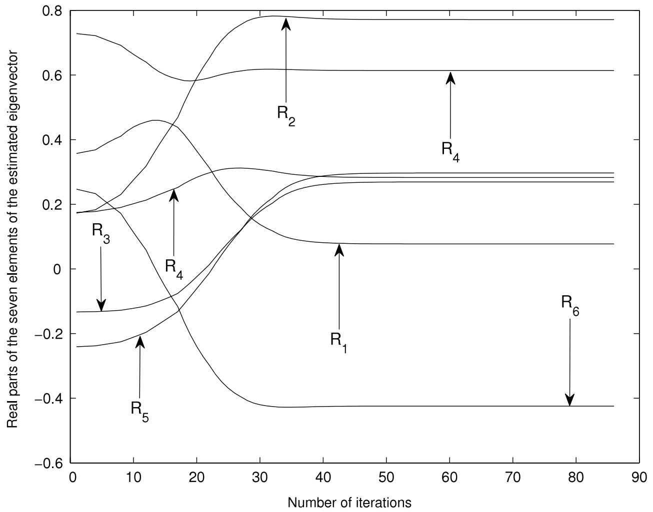

Fig. 5. Transient behavior of the real part of $P^T Qx(t)$, denoted as $R(t)$, where $x(t)$ is the solution of (15) with initial $x(0)$ as in (16). $R(t)$ should converge to the real part of an eigenvector of A corresponding to the eigenvalue $\mathrm{i}\lambda_7$.

After convergence, we can see that $\gamma(t) \to \lambda_7$ and the estimated vector is equal to

$$\text{the estimated for } u_7 = \begin{pmatrix} 0.0774 \\ 0.7713 \\ 0.2692 \\ 0.2829 \\ 0.2969 \\ -0.4242 \\ 0.6137 \end{pmatrix} = 1.1824 \cdot u_7,$$

which is a constant multiple of u_7. Hence, the estimated vector is an eigenvector of A corresponding to the eigenvalue $\mathrm{i}\lambda_7$.

4. Conclusion

In this paper, we extend the neural network based approaches for computing the largest or smallest eigenvalues and the corresponding eigenvectors of real symmetric matrices to the case of real skew-symmetric matrices by employing Householder reflector and similarity transformations. Given any n-by-n real skew-symmetric matrix, unlike the previous neural network based methods that were summarized by $2n$-dimensional ODEs in [14,15], our proposed approach can be represented by such n-dimensional ODEs as (14) and (15). Our proposed method is sure to reduce the scale of networks due

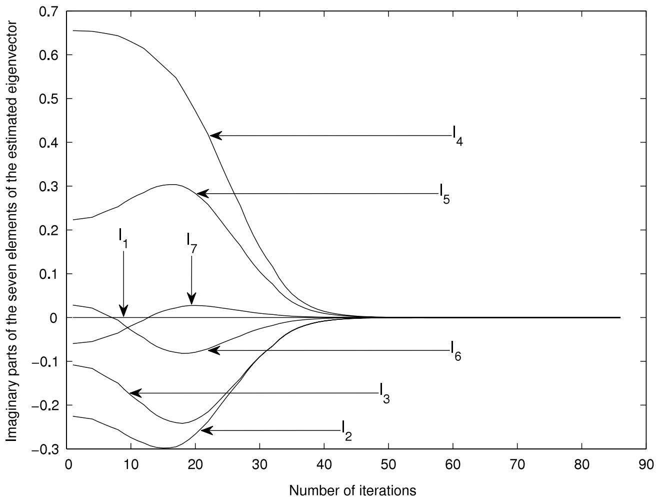

Fig. 6. Transient behavior of the imaginary part of $P^T Qx(t)$, denoted as $I(t)$, where $x(t)$ is the solution of (15) with initial $x(0)$ as in (16). $I(t)$ should converge to the imaginary part of an eigenvector of A corresponding to the eigenvalue $i\lambda_7$.

to its lower dimensionality. Hence, it is more accessible and attractive in practice. Finally, simulations show the validity of our proposed method.

References

[1] C. Chatterjee, V.P. Roychowdhury, M.D. Zoltowski, J. Ramos, Self-organizing and adaptive algorithms for generalized eigendecomposition, IEEE Transactions on Neural Networks 8 (1997) 1518–1530.

[2] E. Oja, J. Karhunen, On stochastic approximation of the eigenvectors and eigenvalues of the expectation of a random matrix, Journal of Mathematical Analysis and Applications 106 (1985) 69–84.

[3] E. Oja, Neural networks, principal components, and subspaces, International Journal of Neural Systems 1 (1989) 61–68.

[4] Y.N. Rao, J.C. Principe, An RLS type algorithm for generalized eigendecomposition, in: IEEE Workshop on Neural Networks for Signal Processing XI, vol. 7, 2001, pp. 263–272.

[5] Y. Xia, An extended projection neural network for constrained optimization, Neural Computation 16 (2004) 863–883.

[6] Y. Xia, J. Wang, A recurrent neural network for solving nonlinear programming convex problems subject to linear constraints, IEEE Transactions on Neural Networks 16 (2005) 379–385.

[7] L. Xu, E. Oja, C. Suen, Modified Hebbian learning for curve and surface fitting, Neural Networks 5 (1992) 441–457.

[8] Q.F. Zhang, Y.W. Leung, A class of learning algorithms for principal component analysis and minor component analysis, IEEE Transactions on Neural Networks 11 (2000) 200–204.

[9] Y. Zhang, F. Yan, H.J. Tang, Neural networks based approach for computing eigenvectors and eigenvalues of symmetric matrix, Computers & Mathematics with Applications 47 (2004) 1155–1164.

[10] J.C. Lv, Y. Zhang, K.K. Tan, Global convergence of Oja's PCA learning algorithm with a non-zero-approaching adaptive learning rate, Theoretical Computer Science 367 (2006) 286–307.

[11] J.S. Taylor, C.K. Williams, N. Cristianini, J.S. Kandola, On the eigenspectrum of the Gram matrix and the generalization error of kernel-PCA, IEEE Transactions on Information Theory 51 (2005) 2510–2522.

[12] Y. Liu, Z.S. You, L.P. Cao, A functional neural network computing some eigenvalues and eigenvectors of a special real matrix, Neural Networks 18 (2005) 1293–1300.

[13] Y. Liu, Z.S. You, L.P. Cao, A recurrent neural network computing the largest imaginary or real part of eigenvalues of real matrices, Computers & Mathematics with Applications 53 (2007) 41–53.

[14] Y. Liu, Z.S. You, L.P. Cao, A functional neural network for computing the largest modulus eigenvalues and their corresponding eigenvectors of an anti-symmetric matrix, Neurocomputing 67 (2005) 384–397.

[15] Y. Liu, Z.S. You, 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 367 (2006) 273–285.

[16] J. Karhunen, J. Joutsensalo, Representation and separation of signals using nonlinear PCA type learning, Neural Networks 7 (1994) 113–128.

[17] K.I. Diamantaras, S.Y. Kung, Principal Component Neural Networks: Theory and Applications, Wiley, New York, 1996.

[18] E. Oja, Principal components, minor components, and linear neural networks, Neural Networks 5 (1992) 927–935.

[19] A. Baker, Matrix Groups: An Introduction to Lie Group Theory, Springer, 2001.

[20] M.D. Plumbley, Geometrical methods for non-negative ICA: manifolds, Lie groups, toral subalgebras, Neurocomputing 67 (2005) 161–197.

[21] S. Fiori, A theory for learning based on rigid bodies dynamics, IEEE Transactions on Neural Networks 13 (2002) 521–531.

[22] A. Householder, Unitary triangularization of a nonsymmetric matrix, Journal of the ACM 5 (1958) 339–342.

[23] G.H. Gloub, C.F. VanLoan, Matrix Computations, Johns Hopkins University Press, Baltimore, MD, 1983.

[24] J.P. Delmas, F. Alberge, Asymptotic performance analysis of subspace adaptive algorithms introduced in the neural network literature, IEEE Transactions on Signal Processing 46 (1998) 170–182.