ISSN 1009－3095 *Journal of Zhejiang University SCIENCE* V.3,No.3,P.298－304,July－Aug.,2002
http://www.periodicals.com.cn; http://www.zju.edu.cn/English

http://www.zjupress.com; http://lib.zju.edu.cn/eindex.htm; jzu_s@mail.hz.zj.cn

Principal component analysis using neural network*

YANG Jian-gang(杨建刚)†, SUN Bin-qiang(孙斌强)

(*Department of Computer Science & Engineering, Zhejiang University, Hangzhou 310027, China*)

†E-mail: yangjg@cs.zju.edu.cn

Received Nov.20,2001; revision accepted Jan.18,2002

Abstract: The authors present their analysis of the differential equation $dX(t)/dt = AX(t) - X^T(t)BX(t)X(t)$, where A is an unsymmetrical real matrix, B is a positive definite symmetric real matrix, $X \in \boldsymbol{R}^n$; showing that the equation characterizes a class of continuous type full-feedback artificial neural network; We give the analytic expression of the solution; discuss its asymptotic behavior; and finally present the result showing that, in almost all cases, one and only one of following cases is true. 1. For any initial value $X_0 \in \boldsymbol{R}^n$, the solution approximates asymptotically to zero vector. In this case, the real part of each eigenvalue of A is non-positive. 2. For any initial value X_0 outside a proper subspace of $\boldsymbol{R}^n$, the solution approximates asymptotically to a nontrivial constant vector $\tilde{Y}(X_0)$. In this case, the eigenvalue of A with maximal real part is the positive number $\lambda = \| \tilde{Y}(X_0) \|_B^2$ and $\tilde{Y}(X_0)$ is the corresponding eigenvector. 3. For any initial value X_0 outside a proper subspace of $\boldsymbol{R}^n$, the solution approximates asymptotically to a non-constant periodic function $\tilde{Y}(X_0, t)$. Then the eigenvalues of A with maximal real part is a pair of conjugate complex numbers which can be computed.

Key words: PCA, Unsymmetrical real matrix, Eigenvalue, Eigenvector, Neural network

INTRODUCTION

It is now well known that principal component analysis (PCA) can be realized in various ways. A PCA network is a one-layer feed forward neural network that can extract the principal components of a stream of input vectors. Typically Hebbian type learning rules are used based on the one-unit learning algorithm originally proposed by Oja (Oja, 1982; 1989). Many different versions and extensions of this basic algorithm have been proposed during recent years (Chen et al., 1998; Lei, 1993; Wang et al., 1996; Yang et al., 2000, Zhang et al., 2000). But all the versions and extensions had the limitation that the matrix for extraction must be symmetric. The new algorithm presented in this paper does not have the limitation.

Let A be an unsymmetrical real matrix. We shall use neural network dynamic system to compute the eigenvalue of A with maximal real part and its corresponding eigenvector. The problem had been investigated for the case when A is symmetric. The corresponding works are presented in many papers.

We consider the following neural network dynamic system:

$$dX(t)/dt = AX(t) - X^T(t)BX(t)X(t) \tag{1}$$

where B is a positive definite symmetric real matrix. $X \in \boldsymbol{R}^n$ may be viewed as the status of neural network. Then Eq.(1) characterizes a class of continuous type full-feedback artificial neural network. It is known from non-linear circuitry that Eq.(1) may be simulated (Luo et al., 1994; 1999). Since this network has the power of parallel computation, continuous dynamics and global network connection, it has high-speed computation ability.

We shall give the analytic expression of the solution of Eq.(1) in Section 2 and discuss its asymptotic behavior in Section 3. The method of computing the eigenvalues and its

* Project (No 69775013) supported by the National Natural Science Foundation of China

eigenvectors are presented in Section 4. Finally we give an example in Section 5.

Notation:

$< u, v >_B = \bar{u}^{\mathrm{T}} B v$, $\| u \|_B = \sqrt{< u, u >_B}$, $\| u \| = \sqrt{\bar{u}^{\mathrm{T}} u}$.

Also "deg" means the degree of a polynomial, and "Re" and "I_m" represent the real part and imaginary part of a complex number.

ANALYTIC EXPRESSION OF THE SOLUTION

In order to give an analytic expression of the solution of the Eq. (1), we shall first choose a suitable basis of $\boldsymbol{C}^n$ in the following lemma.

Lemma 1 There is a basis $\{v_{\alpha,i}\}_{\alpha \in \Lambda, 1 \leqslant i \leqslant n_\alpha}$ of $\boldsymbol{C}^n$ and a collection $\{\lambda_\alpha\}_{\alpha \in \Lambda}$ of complex numbers such that:

$$A v_{\alpha,i} = v_{\alpha,i-1} + \lambda_\alpha v_{\alpha,i}, \quad \alpha \in \Lambda, i = 1, 2, \cdots, n_\alpha \tag{2}$$

where, for convenience, it is assumed that

$$v_{\alpha,0} = 0, \quad \alpha \in \Lambda \tag{3}$$

Proof By the theorem of Jordan canonical form (Stephen et al., 1989), there is an $n \times n$ nonsingular complex matrix S such that

$$S^{-1} A S = J \tag{4}$$

where J is the Jordan's canonical form of A. We put J into the diagonal form:

$$J = \begin{bmatrix} J_1 & 0 & 0 & \cdots & 0 \\ 0 & J_2 & 0 & \cdots & 0 \\ \cdots & \cdots & J_\alpha & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & J_k \end{bmatrix} \tag{5}$$

$$\text{where } J_\alpha = \begin{bmatrix} \lambda_\alpha & 1 & 0 & \cdots & 0 \\ 0 & \lambda_\alpha & 0 & \cdots & 0 \\ & & \cdots\cdots & & \\ 0 & 0 & 0 & \cdots & \lambda_\alpha \end{bmatrix} \tag{6}$$

is an $n_\alpha \times n_\alpha$ Jordan block. Let $\Lambda = \{1, 2, \cdots, k\}$ and $u_{\alpha,i} = e_{n_1 + n_2 + \cdots + n_{\alpha-1} + i}$, $\alpha \in \Lambda$, $i = 1, 2, \cdots, n_\alpha$, where e_k is the column vector of $\boldsymbol{C}^n$ whose entries are 1 on the k-th component and 0 on others, $k = 1, 2, \cdots, n$. By Eq. (5) and Eq. (6),

$$J u_{\alpha,j} = u_{\alpha,i-1} + \lambda_\alpha u_{\alpha,i}, \quad \alpha \in \Lambda, i = 1, 2, \cdots, n_\alpha \tag{7}$$

where $u_{\alpha,0} = 0$, $\alpha \in \Lambda$.

Let $v_{\alpha,i} = S u_{\alpha,i}$, $\alpha \in \Lambda$, $i = 1, 2, \cdots, n_\alpha$ (8)

Since $\{u_{\alpha,i}\}_{\alpha \in \Lambda, 1 \leqslant i \leqslant n_\alpha}$ is a basis of $\boldsymbol{C}^n$ and S is a nonsingular matrix, $\{v_{\alpha,i}\}_{\alpha \in \Lambda, 1 \leqslant i \leqslant n_\alpha}$ is a basis of $\boldsymbol{C}^n$ as well. By Eq. (4),

$$AS = SJ \tag{9}$$

By Eqs. (7), (8) and (9) we have $A v_{\alpha,i} = S u_{\alpha,i-1} + \lambda_\alpha (S u_{\alpha,i} = v_{\alpha,i-1} + \lambda_\alpha v_{\alpha,i}$, $\alpha \in \Lambda$, $i = 1, 2, \cdots, n_\alpha$. The proof is completed.

Theorem 1 For any $X_0 \in \boldsymbol{R}^n$, there is a unique $\boldsymbol{R}^n$-valued function $X(t)$ defined on $\boldsymbol{R}$ satisfying Eq. (1) and

$$X(0) = X_0 \tag{10}$$

Furthermore if we write

$$X_0 = \sum_{\substack{\alpha \in \Lambda \\ 1 \leqslant i \leqslant n}} d_{\alpha,i} v_{\alpha,i} \tag{11}$$

where $d_{\alpha,i} \in \boldsymbol{C}$, $a \in \Lambda$, $i = 1, 2, \cdots, n_\alpha$ and denote $\Lambda_1 = \{\alpha \in \Lambda: \exists\, i \in \{1, 2, \cdots, n_\alpha\}$ takes $d_{\alpha,i} \neq 0\}$ and $m_\alpha = \max\{i \in \{1, 2, \cdots, n_\alpha\}: d_{\alpha,i} \neq 0\}$, $\alpha \in \Lambda_1$, then, the solution may be written as

$$X(t) = \sum_{\substack{\alpha \in \Lambda_1 \\ 1 \leqslant i \leqslant m_\alpha}} P_{\alpha,i}(t) e^{\lambda_\alpha t} \cdot \left(1 + 2\int_0^t \Big\| \sum_{\substack{\beta \in \Lambda_1 \\ 1 \leqslant j \leqslant m_\beta}} P_{\beta,j}(\theta) e^{\lambda_\beta \theta} v_{\beta,j} \Big\|_B^2 d\theta\right)^{-\frac{1}{2}} v_{\alpha,i} \tag{12}$$

where $P_{\alpha,i} = \sum_{r=0}^{m_\alpha - i} (d_{\alpha,i+r}/r!) \cdot t^r$, $\alpha \in \Lambda$, $i = 1, 2, \cdots, m_\alpha$.

To prove theorem 1, we first give the following lemma:

Lemma 2 Assume the functions $f_1(t)$, $f_2(t)$, $\cdots$, $f_k(t)$, $f_{k+1}(t)$ and $h(t)$ satisfy:

$$df_i(t)/dt = f_{i+1}(t) + h(t) f_i(t) \tag{13}$$

$$\text{and} \quad f_i(0) \equiv d_i, \quad i = 1, 2, \cdots, k \tag{14}$$

where d_i, $i = 1, 2, \cdots, k$ are complex constants and

$$f_{k+1} \equiv 0 \tag{15}$$

Then we have $f_i(t) \equiv 0$, $i = 1, 2, \cdots, k$, in the case of $d_1 = d_2 = \cdots = d_k = 0$, and otherwise

$$f_i(t)=\begin{cases}0 & \text{if } i>m\\ p_i(t)e^{\int_0^t h(\theta)d\theta} & \text{if } 1\leqslant i\leqslant m\end{cases} \tag{16}$$

where $m = \max\{i\in\{1,2,\cdots,k\}: d_i\neq 0\}$,

$$P_i(t) = \sum_{r=0}^{m-i}(d_{i+r}/r!)\cdot t^r,\ i = 1,2,\cdots,m. \tag{17}$$

Proof of Theorem 1 Assume $X(t)$ is a function on $\boldsymbol{R}$ with values in $\boldsymbol{C}^n$ which satisfies Eqs. (1) and (10). Taking their conjugate on both side of Eqs. (1) and (10), we conclude that $\overline{X(t)}$ also satisfies Eqs. (1) and (10). From the uniqueness of the solution of an ODE with an initial value (See David Betounes, Differential equations: theory and application: with Maple, Springer-Verlag New York, 2001, p. 89.), it follows that $X(t)=\overline{X(t)}$, that is, $X(t)\in\boldsymbol{R}^n$, $t\in\boldsymbol{R}$. Hence

$$X^{\mathrm{T}}BX(t)=\overline{X(t)}^{\mathrm{T}}BX(t)=\|X(t)\|_B^2 \tag{18}$$

Assume that

$$X(t) = \sum_{\substack{\alpha\in\Lambda\\1\leqslant i\leqslant n_\alpha}} f_{\alpha,i}(t)v_{\alpha,i}, \tag{19}$$

where $f_{\alpha,i}(t)$ is a complex-valued function. By Eqs. (1), (2), (10), (11), (18) and (19), we have $f_{\alpha,i}(0)=d_{\alpha,i}$, $\alpha\in\Lambda$, $i=1,2,\cdots,n_\alpha$ and

$$\sum_{\substack{\alpha\in\Lambda\\1\leqslant i\leqslant n_\alpha}}\left(\frac{\mathrm{d}f_{\alpha,i}(t)}{\mathrm{d}t}\right)v_{\alpha,i}=\sum_{\substack{\alpha\in\Lambda\\1\leqslant i\leqslant n_\alpha}} f_{\alpha,i}(t)(v_{\alpha,i-1}+\lambda_\alpha v_{\alpha,i}) - \|X(t)\|_B^2\sum_{\substack{\alpha\in\Lambda\\1\leqslant i\leqslant n_\alpha}} f_{\alpha,i}(t)v_{\alpha,i} = \sum_{\substack{\alpha\in\Lambda\\1\leqslant i\leqslant n_\alpha}}(f_{\alpha,i+1}(t) + (\lambda_\alpha - \|X(t)\|_B^2)f_{\alpha,i}(t))v_{\alpha,i} \tag{20}$$

where $f_{\alpha,n_\alpha+1}\equiv 0$, $\alpha\in\Lambda$. Comparing coefficients of $v_{\alpha,i}$ on both sides of Eq. (20), we have

$$\frac{\mathrm{d}f_i(t)}{\mathrm{d}t}=f_{i+1}(t)+(\lambda_\alpha-\|X(t)\|_B^2)f_i(t).$$

By Lemma 2, when $\alpha\in\Lambda\setminus\Lambda_1$, $f_{\alpha,i}(t)\equiv 0$, $i=1,2,\cdots,n_\alpha$, and when $\alpha\in\Lambda_1$,

$$f_{\alpha,i}(t)=\begin{cases}0 & \text{if } i>m_\alpha\\ p_{\alpha,i}(t)e^{\int_0^t(\lambda_\alpha-\|X(\theta)\|_B^2)d\theta} & \text{if } 1\leqslant i\leqslant m\end{cases}$$

Define a positive function $E(t) = \mathrm{e}^{-\int_0^t\|X(\theta)\|_B^2)d\theta}$. Thus we have

$$\begin{aligned}X(t) &= \sum_{\substack{\alpha\in\Lambda\\1\leqslant i\leqslant n_\alpha}} f_{\alpha,i}(t)v_{\alpha,i} = \sum_{\substack{\alpha\in\Lambda_1\\1\leqslant i\leqslant n_\alpha}} f_{\alpha,i}(t)v_{\alpha,i}\\ &= \sum_{\substack{\alpha\in\Lambda_1\\1\leqslant i\leqslant n_\alpha}} p_{\alpha,i}(t)\mathrm{e}^{\int_0^t(\lambda_\alpha-\|X(\theta)\|_B^2)\mathrm{d}\theta}v_{\alpha,i}\\ &= E(t)\cdot\sum_{\substack{\alpha\in\Lambda_1\\1\leqslant i\leqslant n_\alpha}} p_{\alpha,i}(t)\mathrm{e}^{\lambda_\alpha t}v_{\alpha,i}\end{aligned}$$

and it follows that, $\frac{\mathrm{d}(E(t))^{-2}}{\mathrm{d}t} = 2\cdot\left\|\sum_{\substack{\alpha\in\Lambda_1\\1\leqslant i\leqslant m_\alpha}} p_{\alpha,i}(t)\mathrm{e}^{\lambda_\alpha t}v_{\alpha,i}\right\|_B^2$. But $E(0) = 1$. Therefore we get

$$E(t) = \left(1+2\int_0^t\left\|\sum_{\substack{\alpha\in\Lambda_1\\1\leqslant i\leqslant m_\alpha}} p_{\alpha,i}(\theta)\mathrm{e}^{\lambda_\alpha\theta}v_{\alpha,i}\right\|_B^2\mathrm{d}\theta\right)^{-\frac{1}{2}} \tag{21}$$

Now, we can conclude Eq. (12) from Eqs. (20) and (21). It is not difficult to verify that Eq. (12) actually satisfies Eqs. (1) and (10). That completes the proof of theorem 1.

THE ASYMPTOTIC BEHAVIOR OF THE SOLUTION

In this section, we analyze the asymptotic behavior of $X(t)$, the solution of the Eq. (1). First we give the following three lemmas.

Lemma 3 Suppose that $P(t)$ is a polynomial with complex coefficients and $Q(t)$ is a nonzero polynomial with complex coefficients. If $\deg P(t)<\deg Q(t)$, then $\lim_{t\to+\infty}P(t)/Q(t)=0$. If $\deg P(t)=\deg Q(t)$, then $\lim_{t\to+\infty}|P(t)/Q(t)-a/b|=0$ where a and b are the leading coefficients of $P(t)$ and $Q(t)$, respectively.

Lemma 4 If λ is a negative real number, and $P(t)$, $Q(t)$ are both nonzero polynomials with complex coefficients, then $\lim_{t\to+\infty}|e^{\lambda t}\cdot(P(t)/Q(t))|=0$.

Lemma 5 Let λ be a positive real number, μ be a complex number, and $P(t)$, $Q(t)$ be nonzero polynomials with complex coefficients whose leading coefficients are a and b, respectively. Then

(1) when either $Re\ \mu > \lambda$, or $Re\ \mu = \lambda$ and $\deg Q(t) > \deg P(t)$ then

$$\lim_{t\to+\infty} \left| P(t)e^{\lambda t} / \int_0^t Q(\theta)e^{\mu\theta}d\theta \right| = 0 .$$

(2) when either $Re\ \mu < \lambda$, or $Re\ \mu = \lambda$ and $\deg Q(t) > \deg P(t)$ then

$$\lim_{t\to+\infty} \left| \int_0^t Q(\theta)e^{\mu\theta}d\theta / P(t)e^{\lambda t} \right| = 0 .$$

(3) when $Re\ \mu = \lambda$ and $\deg Q(t) > \deg P(t)$ then

$$\lim_{t\to+\infty} \left| \int_0^t Q(\theta)e^{\mu\theta}d\theta / P(t)e^{\lambda t} - b/\mu a \cdot e^{(\mu-\lambda)t} \right| = 0$$

Theorem 2 Use notations in Theorem 1 and denote $\Lambda_2 = \{\alpha \in \Lambda_1 : Re\ \lambda_\alpha = \max\{Re\ \lambda_\alpha : \beta \in \Lambda_1\} > 0\}$. If Λ_2 is nonempty, denote $\Lambda_3 = \{\alpha \in \Lambda_2 : m_\alpha = \{\max\ m_\beta : \beta \in \Lambda_2\}\}$. Set $F(t) = \sum_{\substack{\beta\in\Lambda_3\\ \gamma\in\Lambda_3}} 2/(\overline{\lambda_\beta} + \lambda_\gamma) < E_\beta(t), E_\gamma(t) >_B$, where $E_\alpha(t) = d_{\alpha,m_\alpha} e^{\sqrt{-1}\mathrm{Im}\lambda_\alpha t} v_{\alpha,1}$, $\alpha \in \Lambda_3$. Then for any $t \in \boldsymbol{R}$, $F(t)$ is a positive real number and

$$\lim_{t\to+\infty} \| X(t) - Y(t) \| = 0 \qquad (22)$$

where

$$Y(t) = \begin{cases} 0 & \text{if } \Lambda_2 = \varnothing \\ (F(t))^{-\frac{1}{2}} \sum_{\alpha\in\Lambda_3} E_\alpha(t) & \text{if } \Lambda_2 \neq \varnothing \end{cases}$$

Proof Let $f_{\alpha,i}(t) = p_{\alpha,i}(t)e^{\lambda_\alpha t}(1 + 2\int_0^t \| \sum_{\substack{\beta\in\Lambda_1\\ 1\leqslant j\leqslant m_\beta}} p_{\beta,j}(\theta)e^{\lambda_\beta \theta} v_{\beta,j} \|_B^2 d\theta)^{-\frac{1}{2}}, \alpha \in \Lambda_1$, or $1 \leqslant i \leqslant m_\alpha$. If Λ_2 is nonempty, let $R = \max\{Re\lambda_\alpha : \alpha \in \Lambda_1\}$, $m = \max\{m_\alpha : \alpha \in \Lambda_3\}$,

$$G(t) = t^{m-1}/(m-1)! \cdot e^{Rt} \cdot \left[1 + 2\int_0^t \left\| \sum_{\substack{\beta\in\Lambda_1\\ 1\leqslant j\leqslant m}} p_{\beta,i}(\theta)e^{\lambda_\beta \theta} v_{\beta,j} \right\|_B^2 d\theta \right]^{-\frac{1}{2}} .$$

Then R is a positive number and m is a positive integer. We introduce a new norm $|||\cdot|||$ on $\boldsymbol{C}^n$ $||| \sum_{\substack{\alpha\in\Lambda\\ 1\leqslant i\leqslant n_\alpha}} a_{\alpha,i} v_{\alpha,i} ||| = (\sum_{\substack{\alpha\in\Lambda\\ 1\leqslant i\leqslant n_\alpha}} | a_{\alpha,i} |^2)^{\frac{1}{2}}$ where $a_{\alpha,i} \in \boldsymbol{C}$, $\alpha \in \Lambda$, $1 \leqslant i \leqslant n_\alpha$. Since any two norms on $\boldsymbol{C}^n$ are equivalent to each other (Bryan et al., 2000), there are positive numbers C_1, C_2 such that: for any $v \in \boldsymbol{C}^n$, $C_1 ||| v ||| \leqslant \| v \|_B \leqslant C_2 ||| v |||$. Now, using Lemma 3, 4, 5, we can prove the theorem in 4 steps.

Step 1: To prove that $\lim_{t\to+\infty} | f_{\alpha,i}(t) | = 0$ when $\alpha \in \Lambda_1 \setminus \Lambda_3$ or $1 < i \leqslant m_\alpha$.

Step 2: Assuming Λ_2 is nonempty, we can show that

$$\lim_{t\to+\infty} | (G(t))^{-2} - F(t) | = 0. \qquad (23)$$

Step 3: Assuming Λ_2 is nonempty, we can prove that there exist positive constants C_3, C_4 such that for any $t \in \boldsymbol{R}$,

$$C_3 \leqslant F(t) \leqslant C_4. \qquad (24)$$

Step 4: We will prove Eq. (22). We know from Eq. (23) that when t is large enough, $| (G(t))^{-2} - F(t) | \leqslant C_3/2$. Then by Eq. (24), we know that when t is large enough,

$$C_3/2 \leqslant (G(t))^{-2} \leqslant C_4 + C_3/2. \qquad (25)$$

For any $\alpha \in \Lambda_3$, by Lemma 3,

$$\lim_{t\to+\infty} \left| \frac{p_{\alpha,1}(t)}{d_{\alpha,m_\alpha} t^{m-1}/(m-1)!} - 1 \right| = \lim_{t\to+\infty} \left| \frac{\sum_{r=0}^{m-1} d_{\alpha,1+r}/r! \cdot t^r}{d_{\alpha,m_\alpha} t^{m-1}/(m-1)!} - 1 \right| = 0 \qquad (26)$$

Thus it follows from Eqs. (26), (24), (25) and (23) that when $\alpha \in \Lambda_3$ and t is large enough, $| f_{\alpha,1}(t) - (F(t))^{-\frac{1}{2}} e^{\sqrt{-1}\mathrm{Im}\lambda_\alpha t} d_{\alpha,m_\alpha} | \to 0$, *as* $t \to +\infty$. Finally we have

$$\| X(t) - Y(t) \| = \left\| \sum_{\substack{\alpha\in\Lambda_1\\ 1\leqslant i\leqslant m_\alpha}} f_{\alpha,i}(t) v_{\alpha,i} - \sum_{\alpha\in\Lambda_3} (F(t))^{-\frac{1}{2}} e^{\sqrt{-1}\mathrm{Im}\lambda_\alpha t} d_{\alpha,m_\alpha} v_{\alpha,1} \right\| \leqslant$$
$$\sum_{\substack{\alpha\in\Lambda_1\setminus\Lambda_3\\ 1\leqslant i\leqslant m_\alpha}} | f_{\alpha,i}(t) | \cdot \| v_{\alpha,i} \| + \sum_{\substack{\alpha\in\Lambda_3\\ 1\leqslant i\leqslant m_\alpha}} | f_{\alpha,i}(t) | \cdot \| v_{\alpha,i} \| +$$

$$\sum_{\alpha\in\Lambda_3} | f_{\alpha,1}(t) - F(t))^{-\frac{1}{2}} e^{\sqrt{-1}\mathrm{Im}\lambda_\alpha t} d_{\alpha,m_\alpha} | \cdot \| v_{\alpha,1} \| \to 0$$

as $t \to +\infty$. The proof of Theorem 2 is done.

PRINCIPAL COMPONENT EXTRACTION OF MATRIX

We now consider the function $Y(t)$. Assuming Λ_2 is nonempty, set $K_3 = \{\lambda_\alpha : \alpha \in \Lambda_3\}$ and $\Lambda_{3,\lambda} = \{\alpha \in \Lambda_3 : \lambda_\alpha = \lambda\}$ for $\lambda \in K_3$. Then

$$Y(t) = \Big(\sum_{\substack{\mu\in K_3\\ v\in K_3}} 2e^{\sqrt{-1}(\mathrm{Im}\bar{\mu}+\mathrm{Im}v)t}/(\bar{\mu} + v) \cdot < u_\mu, u_v >_B \Big)^{-\frac{1}{2}} \sum_{\lambda\in K_3} e^{\sqrt{-1}\mathrm{Im}\lambda t} u_\lambda \quad (27)$$

where $u_\lambda = \sum_{\alpha\in\Lambda_{3,\lambda}} d_{\alpha,m_\alpha} v_{\alpha,1}$. It is seen from Lemma 1 that u_λ is an eigenvector of A with respect to $\lambda \in K_3$. Since A and X_0 are real, we know that $\{\bar{\lambda} : \lambda \in K_3\} = K_3$ and , $u_{\bar{\lambda}} = \overline{u_\lambda}$, $\lambda \in K_3$. Also from Eq. (27) it follows that

$$\overline{Y(t)} = Y(t). \text{ Thus } Y(t) \in \boldsymbol{R}^n,\ t \in \boldsymbol{R} \quad (28)$$

Denote by V_C and V_R the complex and real subspaces spanned by $\{Y(t) : t \in \boldsymbol{R}\}$ respectively. It is known from Eqs. (27) and (28) that $V_c = \bigoplus_{\lambda\in K_3} Cu_\lambda$, $V_R = V_C \cap \boldsymbol{R}^n$. Thus dim V_R is equal to the number of elements in K_3 and

$$AV_R = A(V_c \cap R^n) \subseteq AV_c \cap AR^n \subseteq (\bigoplus_{\lambda\in K_3} CAu_\lambda) \cap R^n = \bigoplus_{\lambda\in K_3} C(\lambda u_\lambda)) \cap R^n = V_c \cap R^n = V_R \quad (29)$$

that is, V_R is an invariant subspace of A. It is also known that any eigenvector of A in V_C is a multiple of u_λ. Hence its corresponding eigenvalue is in K_3.

We may use the identity

$$Y(t) = \Big(\sum_{\substack{\beta\in\Lambda_3\\ \gamma\in\Lambda_3}} 2 < \widetilde{E}_\beta(t), \widetilde{E}_\gamma(t) >_B / (\overline{\lambda_\beta} + \lambda_\gamma) \Big)^{-\frac{1}{2}} \sum_{\alpha\in\Lambda_3} \widetilde{E}_\alpha(t) \quad (30)$$

where $\widetilde{E}_\alpha(t) = e^{\lambda_\alpha t} d_{\alpha,m_\alpha} v_{\alpha,1}$, $\alpha \in \Lambda_3$ to verify that $Y(t)$ also satisfies the neural network Eq. (1). Before stating Theorem 3, we shall give some notions.

Definition Let $X_1(t)$ and $Y_1(t)$ be $\boldsymbol{R}^n$-valued continuous function on $\boldsymbol{R}$. We say that $X_1(t)$ approximates asymptotically to $Y_1(t)$ if $\lim_{t\to+\infty} \| X_1(t) - Y_1(t) \| = 0$ and $Y_1(t)$ is a periodic function.

Clearly, if $X_1(t)$ approximates asymptotically to both $Y_1(t)$ and $Y_2(t)$, then $Y_1(t) \equiv Y_2(t)$.

It is known from algebraic geometry that there exists a full Leabegue measure densely open subset of $\boldsymbol{R}^{n\times n}$ such that each corresponding $n \times n$ matrix A has different eigenvalues and if λ_1, λ_2 are two eigenvalues of A, then that $Re\lambda_1 = Re\lambda_2$ implies that $\lambda_1 = \lambda_2$ or $\lambda_1 = \overline{\lambda_2}$.

Therefore in the case with noise, A always satisfies the following weaker condition: If λ_1 and λ_2 are the eigenvalues of A with maximal real part, then

$$\lambda_1 = \lambda_2 \text{ or } \lambda_1 = \overline{\lambda_2}. \quad (31)$$

The following lemma is easy to check and left for readers.

Lemma 6 If W is an $n \times 2$ real matrix with rank 2, then the matrix $\boldsymbol{W}^{\mathrm{T}}\boldsymbol{W}$ is invertible.

We shall now give the main theorem of this article.

Theorem 3 Assume that condition Eq. (31) holds. Then one and only one of the following cases is true.

A. For any initial value $X_0 \in \boldsymbol{R}^n$, the solution of Eq. (1) approximates asymptotically to zero vector. In this case, the real part of each eigenvalue of A is non-positive.

B. For any initial value X_0 outside a proper subspace of $\boldsymbol{R}^n$, the solution of Eq. (1) approximates asymptotically to a nontrivial constant vector $\widetilde{Y}(X_0)$. In this case, the eigenvalue of A with maximal real part is the positive number $\lambda = \| \widetilde{Y}(X_0) \|_B^2$ and $\widetilde{Y}(X_0)$ is the corresponding eigenvector.

C. For any initial value X_0 outside a proper subspace of $\boldsymbol{R}^n$, the solution of Eq.(1) approximates asymptotically to a non-constant periodic function $\widetilde{Y}(X_0,t)$. In this case, the smallest positive period T of $\widetilde{Y}(X_0,t)$ is independent of X_0. Take $t_1, t_2 \in \boldsymbol{R}$ with $2(t_1, t_2)/T \notin Z$ and denote the real $n \times 2$ matrix by $W = (\widetilde{Y}(X_0, t_1), \widetilde{Y}(X_0, t_2))$. Then $W^{\mathrm{T}}W$ is invertible. Now we set $(W^{\mathrm{T}}W)^{-1} W^{\mathrm{T}}AW = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then the eigenvalue of A with maximal real part is a pair of conjugate complex numbers $\lambda = a + d + \sqrt{-1}\sqrt{2ad - 4bc - a^2 - d^2}/2$ and $\bar{\lambda}$. $Z = b\widetilde{Y}(X_0, t_1) + (\lambda - a)\widetilde{Y}(X_0, t_2)$ and $\bar{Z}$ are the eigenvectors of A with respect to λ and $\bar{\lambda}$, respectively.

Proof Clearly only one of A, B and C. can hold. Due to Eq.(31), we may only consider the following three cases.

(1) The real part of each eigenvalue of A is non-positive. In this case, for any $X_0 \in \boldsymbol{R}^n$, Λ_2 is empty and $Y(t) \equiv 0$. A holds by Theorem 2.

(2) The eigenvalue of A with maximal real part is a positive number λ.

Denote by U the proper subspace of $\boldsymbol{R}^n$ which is the intersection of $\boldsymbol{R}^n$ and the complex subspace spanned by $\{v_{\alpha,i}: \alpha \in \Lambda, \lambda_\alpha \neq \lambda, 1 \leqslant i \leqslant n_\alpha\}$. Then it is known that for initial value $X_0 \in \boldsymbol{R}^n \setminus U$ we have $K_3 = \{\lambda\}$. Set $\widetilde{Y}(X_0) = \sqrt{\lambda}u_\lambda / \| u_\lambda \|_B$. It follows that $\lambda = \| \widetilde{Y}(X_0) \|_B^2$ and $\widetilde{Y}(X_0)$ is the eigenvector with respect to λ. By Eq.(27), we have $Y(t) \equiv \widetilde{Y}(X_0)$. It is also seen from Theorem 2 that the solution of Eq.(1) with initial value X_0 approximates asymptotically to $\widetilde{Y}(X_0)$. Therefore B holds.

(3) The eigenvalues of A with maximal real part is a pair of conjugate complex numbers λ and $\bar{\lambda}$ with $\mathrm{Im}\lambda > 0$, $Re\,\lambda > 0$.

Denote by U' the proper subspace of $\boldsymbol{R}^n$ which is the intersection $\boldsymbol{R}^n$ and the complex linear subspace spanned by $\{v_{\alpha,i}: \alpha \in \lambda, \lambda_\alpha \neq \lambda, \lambda_\alpha \neq \lambda, \lambda_\alpha \neq \bar{\lambda}, 1 \leqslant i \leqslant n_\alpha\}$. Then when $X_0 \in \boldsymbol{R}^n \setminus U'$, $K_3 = \{\lambda, \bar{\lambda}\}$. Set

$$\widetilde{Y}(X_0, t) = \Big(\frac{2}{Re\lambda} \| u_\lambda \|_B^2 + \frac{1}{\lambda} e^{2\sqrt{-1}\mathrm{Im}\lambda t} < u_{\bar{\lambda}}, u_\lambda >_B + \frac{1}{\bar{\lambda}} e^{2\sqrt{-1}\mathrm{Im}\bar{\lambda} t} < u_\lambda, u_{\bar{\lambda}} >_B \Big)^{\frac{1}{2}}$$

$$\left(e^{\sqrt{-1}\mathrm{Im}\lambda t} u_\lambda + e^{\sqrt{-1}\mathrm{Im}\bar{\lambda} t} u_{\bar{\lambda}} \right) \quad (32)$$

By Eq.(27), we get $Y(t) = \widetilde{Y}(X_0, t)$. It is known from Eq.(32) that the smallest positive period of $\widetilde{Y}(X_0, t)$ is $T = 2\pi/\mathrm{Im}\lambda$ independent of X_0. It follows from theorem 2 that the solution $X(t)$ of Eq.(1) with initial value X_0 approximates asymptotically to $\widetilde{Y}(X_0, t)$. When $t_1, t_2 \in \boldsymbol{R}$ and $2(t_1, t_2)/T \notin Z$, $\widetilde{Y}(X_0, t_1)$ and $\widetilde{Y}(X_0, t_2)$ are linearly independent and form a basis of V_R. Therefore the rank of $W = (\widetilde{Y}(X_0, t_1), \widetilde{Y}(X_0, t_2))$ is 2. By Lemma 6, $W^{\mathrm{T}}W$ is invertible. It is known from Eq.(29) that there exists a real 2×2 matrix $\widetilde{A}$ such that $AW = W\widetilde{A}$. Thus $W^{\mathrm{T}}AW = (W^{\mathrm{T}}W)\widetilde{A}$ and $\widetilde{A} = (W^{\mathrm{T}}W)^{-1} W^{\mathrm{T}}AW = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

Let λ_1 and λ_2 be the eigenvalues of $\widetilde{A}$. Set

$$Z_i = b\widetilde{Y}(X_0, t_1) + (\lambda_i - a)\widetilde{Y}(X_0, t_2), \quad i = 1, 2$$

Since $\lambda_i^2 - (a + d)\lambda_i + (ad - bc) = 0$, $i = 1, 2$, we may easily verify that $AZ_i = \lambda_i Z_i$, $i = 1, 2$. Due to Eq.(30), $\lambda_1, \lambda_2 \in \{\lambda, \bar{\lambda}\}$. We may assume $\lambda_1 = \lambda$, $\lambda_2 = \bar{\lambda}$. Then

$$\lambda = \lambda_1 = (a + d + \sqrt{-1}\sqrt{2ad - 4bc - a^2 - d^2})/2,$$

and $Z = Z_1$, $\bar{Z} = Z_2$. Therefore C. follows. That completes the proof.

EXSAMPLE

Considering a 2×2 real matrix $A = \begin{pmatrix} 2 & 5 \\ -1 & 6 \end{pmatrix}$, and a random initial value $X_0 = \begin{pmatrix} 0.8 \\ 0.5 \end{pmatrix}$, system (1) will evolve asymptotically to a non-constant periodic function $X(t)$.

We take

$$X(t_1) = \begin{pmatrix} -1.91865905 \\ -1.25761778 \end{pmatrix},$$

$$X(t_2) = \begin{pmatrix} -0.81851740 \\ 1.08864150 \end{pmatrix},$$

that follows

$$W = (X(t_1), X(t_2)) = \begin{pmatrix} -1.91865905 & -0.81851740 \\ -1.25761778 & 1.08864150 \end{pmatrix},$$

$$\widetilde{A} = (W^{\mathrm{T}}W)^{-1}W^{\mathrm{T}}AW = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 5.01225265 & 0.62137115 \\ -3.25836731 & 2.98774735 \end{pmatrix},$$

$$\lambda = \frac{a+d+\sqrt{-1}\sqrt{2ad-4bc-a^2-d^2}}{2} = 4+\sqrt{-1}$$

$$Z = bX(t_1)+(\lambda-a)X(t_2) = \begin{bmatrix} 7.08024231 & -0.81851740\sqrt{-1} \\ 2.99580040 & +1.08864150\sqrt{-1} \end{bmatrix},$$

then the eigenvalues of A with maximal real part is a pair of conjugate complex numbers λ and $\bar{\lambda}$, Z and $\bar{Z}$ are the eigenvectors of A with respect to λ and $\bar{\lambda}$, respectively.

CONCLUSIONS

We would like to list two most important cases:

1. The eigenvalue of A with maximal real part is a positive number λ with multiplicity 1. Let v and $-v$ be the eigenvectors of A with respect to λ such that $\|v\|_B = \lambda$. Then $(A-\lambda)(\boldsymbol{R}^n)$ is a hyperplane of $\boldsymbol{R}^n$ which separates the whole space into two half spaces. Any solution of the Eq. (1) with initial value in the open half space containing v always converges to v, while any solution with initial value in the open half space containing $-v$ always converges to $-v$.

2. The eigenvalues of A with maximal positive real part is a pair of complex numbers λ and $\bar{\lambda}$, and the multiplicity of λ is 1. In this case, $\boldsymbol{R}^n = Ker(A^2 - 2\,\mathrm{Re}\,\lambda\cdot A + |\lambda^2|\cdot I)\bigoplus(A^2-2\,\mathrm{Re}\,\lambda\cdot A + |\lambda^2|\cdot I)(\boldsymbol{R}^n)$, and $Ker(A^2-2\,\mathrm{Re}\,\lambda\cdot A+|\lambda^2|\cdot I)$ is a two dimensional space denoted by P. The closed orbit $\{\widetilde{Y}(X_0,t): t\in R\}$ in Theorem 3 does not depend on the initial value X_0 in $\boldsymbol{R}^n \setminus (A^2-2\,\mathrm{Re}\,\lambda\cdot A+|\lambda^2|\cdot I)(\boldsymbol{R}^n)$. $\widetilde{Y}(X_0,t)$ is just the periodic solution of Eq. (1) with the initial value X'_0, where X'_0 is the intersection of the closed obit and the straight line from 0 to the projection of X_0 on P.

Hence, all of the constant solutions in the first case and the periodic solutions in the second case have good stability. In fact, in some sense, most matrices satisfy case (1) or (2), or all the eigenvalues have non-positive real part. In the last case, we can adjust the matrix by adding a suitable positive scalar matrix so that we can apply the method to compute the eigenvalues and eigenvectors.

The method is also suitable for complex matrices with only a little change.

References

Bryan P. Rynne, Martin A. Youngson, 2000. Linear Functional Analysis. Springer.

Chen Tianping, Shun Ichi Amari, Qin Lin, 1998. An unified algorithm for principal and minor components extraction. *Neural Networks* **11**(3): 385−390.

David Betounes, 2001. Differential equations: theory and application. Maple, Springer-Verlag, New York, p. 89.

Lei Xu, 1993. Least MSE Reconstruction: A principle for self-organizing nets, *Neural Networks*, **6**(6): 627−648.

Luo Falong, Li Yanda, Wang Zhe, 1994. Finding eigenvectors of a class of positive matrix by neural networks. *Science in China*, **24**(1): 83−88 (in Chinese).

Luo Falong, Rolf Unbehauen, 1999. Comments on: A unified algorithm for principal and minor components extraction, *Neural Networks* **12**(2): 393.

Oja, E., 1982. A simplified neuron model as a principle component analyzer. *J Math Biol*, **15**: 267−273.

Oja, E., 1989. Neural networks, principal components, and subspace, *Int J Neural System*, **1**: 61−68.

Stephen H. Friedberg, Arndd J. Insel, Lawrence E. Spence, 1989. Linear Algebra. Prentice Hall, Inc.

Wang Zhe, Li Yanda, Luo Falong, 1996, A neural network learning algorithm for PCA and MCA. *Acta Electronica Sinica*, **24**: 12−16 (in Chinese).

Yang Jiangang, Wang Kai, Yang huayong et al., 2000. A real-time adaptive control algorithm using neural nets with perturbation. *Journal of Zhejiang University SCIENCE*, **1**(1): 61−65.

Yang Jiangang, Wang Ruming., 2000. A Rapid Fuzzy Rule Extraction Method For Fuzzy Controller. *Journal of Zhejiang University SCIENCE*, **1**(3): 311−376.

Zhang Yi, Wang Pingan, Zhou Mingtian, 2000, Computing eigenvalues and eigen vectors of matrix by neural networks. *CHINESE J. COMPUTERS*, **23**(1): 71−76 (in Chinese).