Self-Normalizing Dual Systems for Minor and Principal Component Extraction

Mohammed A. Hasan
Department of Electrical & Computer Engineering
University of Minnesota Duluth
E.mail:mhasan@d.umn.edu

Abstract: In this paper classes of globally stable dynamical systems for dual-purpose extraction of principal and minor components are analyzed. The proposed systems may apply to both the standard and the generalized eigenvalue problems. Lyapunov stability theory and LaSalle invariance principle are used to derive invariant sets for these systems. Some of these systems may be viewed as generalizations of known learning rules such as Oja's and Xu's systems and are shown to be applied, with some modifications, to symmetric and nonsymmetric matrices. Numerical examples are provided to examine the convergence behavior of the dual-purpose minor and principal component analyzers.

Keywords: Principal components, minor components, generalized eigenvalue problem, Liapunov stability, global convergence, Oja's learning rule, Rayleigh quotient, dual-purpose MCA/PCA systems.

1 Introduction

Principal (PCA) and minor (MCA) component analyzers of a symmetric matrix are matrix differential equations that converge to the eigenvectors associated with the largest and smallest eigenvalues, respectively. Similarly, principal (PSA) and minor (MSA) subspace analyzers of a symmetric matrix are matrix differential equations that converge to a matrix whose columns's span is the subspace spanned by the eigenvectors corresponding to the largest and smallest eigenvalues, respectively. PCA and MCA are useful tools in adaptive antenna arrays in signal processing, multiuser detection in wireless communication, and truncated model reduction tasks.

After the pioneering work of Oja [1], Sanger [2], Xu [3], Amari [4], and others [5,6], numerous learning rules for principal component analysis have been developed in the literature. Some of these rules are modifications of the original PCA learning systems. As indicated in [6], the task of developing an MCA flow is perceived as being more complicated than that for a PCA flow. The present work shows that perhaps there are as many MCA/MSA dynamical flows as there are PCA/PSA flows.

A common method for converting a PCA/PSA flow into an MCA/MSA one is to change the sign of the given matrix [6], or by using the inverse of the original matrix. However, inverting a large matrix is a costly task, and changing the sign of the original matrix does not always generate a stable system unless frequent orthonormalization is employed during the numerical implementation. In this paper we propose a framework to generate classes of stable dynamical systems that can be easily converted from PCA flow into MCA flow and vice versa.

Throughout this paper, the following notation will be used. The symbols $\mathbb{R}$, and $\mathbb{N}$ denote the set of real numbers, and the set of positive integers, respectively. The transpose of a real matrix x is denoted by x^T, and the derivative of x with respect to time is written as x'. If $B = [b_{ij}]$ is a square matrix of size n, then $tr(B) = \sum_1^n b_{ii}$ denotes the trace of B, and $diag(B)$ is a diagonal matrix with diagonal elements $b_{11}, \cdots, b_{nn}$. The identity matrix of appropriate dimension is expressed with the symbol I. Also, the derivative of $V(x)$ with respect to time along a trajectory $x' = f(x)$ is denoted by $\dot{V}$. Finally, it will be assumed that the matrix A has distinct eigenvalues unless otherwise stated.

2 Mathematical Preliminaries

In this section, we introduce several known results from Lyapunov stability theory of dynamical systems. Let $g(x) : \mathbb{R}^{n\times p} \to \mathbb{R}^{n\times p}$, $p \le n$, be continuously differentiable function and consider the dynamical system

$$x' = g(x). \tag{1}$$

A set $S \in \mathbb{R}^{n\times p}$ is an *invariant set* for the system (1) if every trajectory $x(t)$ which starts from a point in S remains in S for all time. For example, any equilibrium point is an invariant set. The domain of attraction of an equilibrium point is also an invariant set.

We state next a few stability results for nonlinear autonomous systems. The invariant set theorems reflect the intuition that the decrease of a Liapunov function V has to gradually vanish. In other words $\dot{V}$ has to converge to zero because V is lower bounded. Proofs of the results below can be found in [7].

Theorem 1 (Local Invariant Set Theorem) [7]. *Consider the autonomous system (1) with g continuous and let $V(x) : \mathbb{R}^n \to \mathbb{R}$ be a scalar function with continuous first partial derivatives. Assume that*

1. *for some $l > 0$, the set Ω_l defined by $V(x) \le l$ is bounded.*
2. *$V'(x) \le 0$ for all x in Ω_l.*

Let R be the set of all points within Ω_l where $V'(x) = 0$ and M be the largest invariant set in R. Then, every solution $x(t)$ originating in Ω_l tends to M as $t \to \infty$.

In Theorem 1, the word largest means that M is the union of all invariant sets within R. Notice that R is not necessarily connected, nor is the set M.

We state next a well known result about Lagrange stability.

Theorem 2 (A Lagrange Stability Theorem) [7]. *Let W be a bounded neighborhood of the origin and let W^c be its complement (W^c is the set of all points outside W). Assume that $V(x)$ is a scalar function with continuous first partial derivatives in W^c and satisfying:*

1. *$V(x) > 0$ for all $x \in W^c$,*
2. *$\dot{V}(x) \le 0$ for all $x \in W^c$,*
3. *$V(x) \to \infty$ as $||x|| \to \infty$.*

Then each solution of (1) is bounded for all $t > 0$.

3 Dual-Purpose Systems

We develop here a number of dynamical systems that can be converted between PCA/PSA rules and MCA/MSA rules. To prove stability for the original and the converted systems, the following theorem is needed. It examines the trace of a product of symmetric and anti-symmetric matrics.

Proposition 3. *Let $P, S \in \mathbb{R}^{n\times n}$ be symmetric and antisymmetric matrices, respectively. Then $tr(PS) = 0$.*

Proof. Let $P = [p_{ij}]$ and $S = [s_{ij}]$, then

$$tr(PS) = \sum_{i=1}^{n}\sum_{k=1}^{n} p_{ik}s_{ki}.$$

Since $s_{ll} = 0, l = 1, \cdots n$, $p_{ik} = p_{ki}$, and $s_{ki} = -s_{ik}$ it follows that

$$\begin{aligned} tr(PS) &= \sum_{l}^{n} p_{ll}s_{ll} + \sum_{k=1}^{n-1}\sum_{i=k+1}^{n} \{p_{ik}s_{ki} + p_{ki}s_{ik}\} \\ &= \sum_{l}^{n} p_{ll}s_{ll} + \sum_{k=1}^{n-1}\sum_{i=k+1}^{n} \{p_{ik}s_{ki} - p_{ik}s_{ki}\} = 0. \end{aligned}$$

3.1 MSA/PSA Dynamical Systems

In this section, dynamical systems for implementing dual-purpose MCA/MSA and PCA/PSA algorithms are developed. Let Ω_1 be the set $\Omega_1 = \{x : x^T x \text{ is positive definite}\}$. In what follows, it will be assumed that the initial condition x_0 for dynamical systems satisfies $x_0 = x(0) \in \Omega_1$. The derivation of these systems are motivated by the idea that the Rayleigh quotient $tr\{(x^T Ax)(x^T x)^{-1}\}$ is bounded over the set Ω_1 regardless of whether A is positive definite or not. Additionally, it remains bounded even if A is replaced with $-A$. Similarly, the inverse Rayleigh quotient $tr\{(x^T x)(x^T Ax)^{-1}\}$ is bounded over the set $\Omega_2 = \{x : x^T Ax \text{ is positive definite}\}$ provided that A is positive or negative definite. To examine the critical points of the Rayleigh quotient and the inverse Rayleigh quotient for the simple case where $x \in \mathbb{R}^{n\times 1}$, let A be symmetric and x is one dimensional. One can show that the right hand side of each of the systems

$$x' = xx^T Ax - Axx^T x, \qquad (2a)$$
$$x' = Axx^T x - xx^T Ax, \qquad (2b)$$

appears in the numerator of the gradient of the Rayleigh and the inverse Rayleigh quotients. Specifically, the system (2a) can be expressed as

$$x' = \nabla_x(\frac{x^T x}{x^T Ax})(x^T Ax)^2, \qquad (3)$$

while the system (2b) can be written as

$$x' = -\nabla_x(\frac{x^T Ax}{x^T x})(x^T x)^2. \qquad (4)$$

This shows that $xx^T Ax - Axx^T x = \nabla_x(\frac{x^T x}{x^T Ax})(x^T Ax)^2 = -\nabla_x(\frac{x^T Ax}{x^T x})(x^T x)^2$. Thus both systems (2a) and (2b) are gradient-like systems since $(x^T Ax)^2$ and $(x^T x)^2$ are positive definite if $x \in \Omega_1$ and A is positive definite.

To understand the behavior of the Rayleigh quotient along the trajectory of (2a) and (2b) let $f(x) = \frac{x^T Ax}{x^T x}$, $A^T = A$, then $\dot{f} = \nabla_x(f(x))^T x' = -\nabla_x(f(x))^T \nabla_x(f(x))(x^T Ax)^2 \le 0$. Consequently, $f(x(t))$ is decreasing function for $t \ge 0$ and since it is bounded below, $\lim_{t\to\infty} f(x(t))$ exists. Also note that $V(x) = x^T x$ remains constant along the trajectory of the system (2a), while the function $V = x^T Ax$ is decreasing since $\dot{V} = (x^T Ax)^2 - x^T A^2 xx^T x \le 0$. This implies that $x^T(t)Ax(t) \le x_0^T Ax_0$ and $x(t)^T x(t) = x_0^T x_0$ for $t \ge 0$.

From the previous discussion, we state the following theorem.

Theorem 4. *The systems (2a) and (2b) are stable and if $x(t)$ is a solution of either systems for $t \ge 0$, then $x(t)^T x(t) = x(0)^T x(0)$ for each $t \ge 0$.*

The next result provides several generalizations of the systems (2a) and (2b).

Theorem 5. *Consider the dynamical systems*

$$x' = xK(x) - Axx^T x, \qquad (5a)$$
$$x' = Axx^T x - xK(x), \qquad (5b)$$

where $K(x) : \mathbb{R}^{n\times p} \to \mathbb{R}^{p\times p}$, $p \le n$ is a continuously differentiable function. If $K + K^T = x^T Ax + x^T A^T x + \alpha(I - x^T x)B(x)$, where $\alpha \ge 0$ and $B(x) + B(x)^T$ is positive definite, then the systems (5a) and (5b) are stable.

Outline of Proof: By considering a Liapunov function of the form $V(x) = \frac{1}{4}tr((x^T x - I)^2)$, it can be shown that the time derivative of V along the trajectory $x(t)$ of the system (5a) is

$$\begin{aligned} \dot{V} &= tr\{(x^T x - I)(K^T - x^T Ax)x^T x \\ &= \frac{1}{2}tr\{(x^T x - I)(K^T + K - x^T Ax - x^T A^T x)x^T x\} \\ &= \frac{-\alpha}{4}tr\{(x^T x - I)^2(B(x) + B(x)^T)\} \le 0. \end{aligned}$$

Since $V(x) \to \infty$ as $||x|| \to \infty$, Theorem 1 guarantees that the system (5a) is stable. Similarly the system (5b) is stable.

Remark 1: It is interesting to note that Systems (2a), (2b), (5a), (5b) are stable for any matrix A. By carefully examining the proofs of Theorem 4 and 5, one can also show that the systems

$$x' = xx^T A^T x - Axx^T x, \qquad (2a)'$$
$$x' = Axx^T x - xx^T A^T x, \qquad (2b)'$$

are stable. This means that the systems (2a), (2a)', (2b), and (2b)' converge to the minor or the principal subspaces for any matrix A having distinct eigenvalues.

Special Cases: Based on the above theorem, several variations of System (2a) and (2b) may be derived. For example, assume that in Theorem 5 we set $K - x^T Ax = \alpha B(x)(I - x^T x)$, where $B(x) + B(x)^T$ is positive definite, and $\alpha \ge 0$. Then System (5a) simplifies to

$$x' = xx^T Ax - Axx^T x - \alpha xB(x)(x^T x - I). \qquad (6a)$$

Similarly, System (5b) simplifies to

$$x' = Axx^T x - xx^T Ax - \alpha xB(x)(x^T x - I). \qquad (6b)$$

In particular, when $B(x) = I$, then the following MSA/PSA systems are resulted:

$$x' = xx^T Ax - Axx^T x - \alpha x(x^T x - I), \qquad (7a)$$
$$x' = Axx^T x - xx^T Ax - \alpha x(x^T x - I). \qquad (7b)$$

When A is symmetric, other variations follow by incorporating the term $-\alpha A^k x(x^T x - I)$ into Systems (2a) and (2b):

$$x' = \pm\{xx^T Ax - Axx^T x\} - \alpha A^k x(x^T x - I), \qquad (8)$$

where k is any integer and $\alpha \ge 0$. Here, the choices of the + and − signs yield MSA and PSA systems, respectively. The systems described in (8) can be shown to converge provided that A^k is positive definite and $x(0) \in \Omega_1$. It is interesting to note that if the + sign is chosen and $\alpha = k = 1$, we obtain

$$x' = Ax - xx^T Ax \qquad (9)$$

which is one form of Oja's subspace system. When $k = 1$ and $\alpha = 1$, the system (8) with the + sign reduces to

$$x' = xx^T Ax - Ax(2x^T x - I). \qquad (10)$$

This MSA system is known in the literature and is analyzed in [10].

3.1.1 Conversion to MCA/PCA Systems

To convert MSA/PSA systems into MCA/PCA learning rules, we may incorporate a diagonal matrix D in the above systems. For example if in Systems (8) we replace I with D, the following systems are obtained.

$$x' = \pm\{xx^TAx - Axx^Tx\} - \alpha A^k x(x^Tx - D). \qquad (11)$$

If $k = 1$ and $\alpha = 1$, the two systems in (11) simplify to

$$x' = AxD - xx^TAx, \qquad (12a)$$

$$x' = xx^TAx - Ax(2x^Tx - D). \qquad (12b)$$

The system (12a) is sometimes called Xu's weighted PCA rule [3]. Also the system (12b) is an MCA version of the minor subspace system (10).

Remark 2. From the previous discussion, it seems that each of the dual-purpose learning differential equation derived so far can be expressed as $x' = xf(x, A, D) - Axg(x, A, D) + h(x, A, D)$, where D is a diagonal matrix, and f, g and h are chosen so that the function $V(x) = \frac{1}{2}trace(x^Tx))$ has zero time derivative along the trajectory of the differential equation $x' = xf(x, A, D) - Axg(x, A, D)$ and $\dot{V}$ is negative semidefinite along the trajectory of the differential equation $x' = h(x, A, D)$. Thus the dynamical system $x' = -xf(x, A, D) + Axg(x, A, D) + h(x, A, D)$ is also stable and may converge to PSA/PCA. For simplicity of implementation $h(x, A, D)$ is normally chosen as $\alpha A^k x(x^Tx - D)$ for $k = 0, 1$. Many other expressions are considered in the next sections.

3.2 Other Dual PCA/MCA Systems

We now use logarithmic cost functions to derive learning rules that can be used as a PCA/PSA and MCA/MSA dynamical systems by merely switching a sign. Other MCA/PCA learning rules based on logarithmic cost functions are derived in [8,9]. Let $f(x) = \frac{\pm 1}{2}\{trace(\log(x^TAx)) - \frac{1}{2}trace(\log(x^Tx))\}$, where A is positive definite and $x \in \Omega_1$. Then the gradient of f is $\nabla f(x) = \pm\{Ax(x^TAx)^{-1} - x(x^Tx)^{-1}\}$. Thus we obtain the following gradient systems:

$$x' = Ax(x^TAx)^{-1} - x(x^Tx)^{-1}, \qquad (13a)$$

which is a PSA learning rule, and

$$x' = x(x^Tx)^{-1} - Ax(x^TAx)^{-1}, \qquad (13b)$$

which is an MSA learning rule.

We note that if $V(x) = \frac{1}{2}tr(x^Tx)$, then $\dot{V} = 0$ along any trajectory of the systems (13a)-(13b). Consequently, $V(x(t)) = V(x(0))$ for each $t \geq 0$, or equivalently, $x(t)^Tx(t) = x_0^Tx_0$. Although these two dynamical systems are self-normalized, with normalization depending on the initial condition, the main drawback is that only principal or minor subspaces can be obtained. Numerical simulations have shown that MCA/PCA systems will result if a positive definite diagonal matrix D having distinct eigenvalues is incorporated so that:

$$x' = AxD(x^TAx)^{-1} - xD(x^Tx)^{-1}, \qquad (14a)$$

$$x' = xD(x^Tx)^{-1} - AxD(x^TAx)^{-1}. \qquad (14b)$$

By considering the Liapunov function $V(x) = \frac{1}{2}tr(x^Tx)$, it follows that the time derivative is $\dot{V} = 0$ along any trajectory of the systems (14a) and (14b). This implies that (14a) and (14b) are globally stable over the set Ω_1.

To examine the behavior of the limiting solutions, let $P = \lim_{t\to\infty} x(t)^Tx(t)$ and $B = \lim_{t\to\infty} x(t)^TAx(t)$. Then $PDP^{-1} = BDB^{-1}$ or $DP^{-1}B = P^{-1}BD$. Therefore, $P^{-1}B$ is diagonal (see Proposition 6 [8]). Now assume that $P^{-1}B = D_1$ then $B = PD_1 = D_1P$. The second equality follows since P and B are symmetric. The eigenvalues of $P^{-1}B$ are eigenvalues of A. Thus the diagonal elements of D_1 are distinct. This implies that $P = D_2$ for some diagonal matrix D_2 (see Proposition 6 [8]). Now $B = PD_1 = D_1D_2$ is diagonal.

Motivated by the systems (14a) and (14b), the systems (2a) and (2b) are modified analogously by inserting a diagonal matrix D so that

$$x' = xDx^TAx - AxDx^Tx, \qquad (15a)$$

$$x' = AxDx^Tx - xDx^TAx. \qquad (15b)$$

After numerous simulations, it should be pointed out that (15a) and (15b) do not represent MCA or PCA learning rules. They only converge to minor and principal subspaces, respectively.

However, to convert the systems (15a)-(15b) into MCA/PCA systems, one may add a penalty term such as $-\alpha\, x(x^Tx - D)$, where $\alpha > 0$, so that the modified systems are

$$x' = xDx^TAx - AxDx^Tx - \alpha x(x^Tx - D), \qquad (16)$$

$$x' = AxDx^Tx - xDx^TAx - \alpha x(x^Tx - D). \qquad (17)$$

Simulations have shown that these systems have very good convergence behavior to MCA and PCA. To show that theoretically, assume that A is positive definite and D is positive definite diagonal matrices of appropriate dimensions. Let $P = \lim_{t\to\infty} x(t)^Tx(t)$ and $B = \lim_{t\to\infty} x(t)^TAx(t)$. Then Equation (16) implies that $PDB = BDP - \alpha P(P - D)$ or $PDB - BDP = -\alpha P(P - D)$. Taking the transpose of both sides yields $BDP - PDB = -\alpha(P - D)P$. By adding the last two equation we obtain $(P - D)P + P(P - D) = 0$. Since P is positive definite, it follows that $P = D$. Hence $D^2B = BD^2$. Since all eigenvalues of D^2 are distinct, it follows that $B = D_1$ for some diagonal matrix D_1 (see Proposition 6 [8]). This shows that both $\lim_{t\to\infty} x(t)^Tx(t)$ and $\lim_{t\to\infty} x(t)^TAx(t)$ are diagonal. Analogous proof holds for the system (17).

Similar analysis shows that each of the systems

$$x' = xx^TAx - Axx^Tx - \alpha x(x^Tx - D), \qquad (18a)$$

$$x' = xx^Tx - Axx^TAx - \alpha x(x^Tx - ddiag(x^Tx)), \qquad (18b)$$

converges to minor components while each of the systems

$$x' = Axx^Tx - xx^TAx - \alpha x(x^Tx - D), \qquad (18c)$$

$$x' = Axx^Tx - xx^TAx - \alpha x(x^Tx - ddiag(x^Tx)), \qquad (18d)$$

converges to principal components. Clearly, Systems (18a)-(18d) are modifications of those of (2a) and (2b).

By adding different penalty terms, many other dual-purpose MCA/PCA learning systems may be obtained:

$$x' = \pm(xx^TAx - Axx^Tx) - \alpha x(x^Tx)^{-1}(x^Tx - D), \qquad (19a)$$

$$x' = \pm(xx^TAx - Axx^Tx) - \alpha x(x^TAx - D), \qquad (19b)$$

$$x' = \pm(xx^TAx - Axx^Tx) - \alpha Ax(x^Tx - D), \qquad (19c)$$

$$x' = \pm(xx^TAx - Axx^Tx) - \alpha Ax(x^TAx - D), \qquad (19d)$$

$$x' = \pm(xx^TAx - Axx^Tx) - \alpha Ax(x^TAx)^{-1}(x^TAx - D), \quad (19e)$$

where $\alpha > 0$. Stability analysis can be established for these systems based on Proposition 3, Theorems 1 and 2.

4 Generalized Eigenvalue Problem

The PCA/PSA and MCA/MSA learning differential equations of the previous sections may be modified to obtain PCA/PSA and MCA/MSA learning differential equations for the generalized eigenvalue problem involving two matrices A and B where B is positive definite and $A^T = A$. Thus we consider the following systems:

$$x' = \pm(Axx^TBx - Bxx^TAx) - \alpha x(x^TBx - D)^m, \qquad (20a)$$

$$x' = \pm(Axx^TBx - Bxx^TAx) - \alpha Ax(x^TBx - D)^m, \qquad (20b)$$

$$x' = \pm(Axx^T Bx - Bxx^T Ax) - \alpha Bx(x^T Bx - D)^m, \quad (20c)$$

$$x' = \pm(B^{-1}Axx^T Bx - xx^T Ax) - \alpha x(x^T Bx - D)^m. \quad (20d)$$

Here $x \in \mathbb{R}^{n\times p}$, D is positive definite diagonal matrix, $\alpha \geq 0$, and $m \geq 0$ is an integer.

In what follows we show that System (20a) is stable. The stability of other systems (20b)-(20d) may be established similarly. Let $s \in \mathbb{N}$ so that $s - 1 + m$ is even and consider $V(x) = \frac{1}{2s} tr((x^T Bx - D)^s)$, where D is a diagonal matrix. Now System (20a) is stable if the following system

$$x' = \pm(B^{-1}Axx^T Bx - xx^T Ax) - \alpha B^{-1}x(x^T Bx - D)^m. \quad (20e)$$

is stable. The time derivative of V along the trajectory $x(t)$ of (20e) is

$$\begin{aligned}\dot{V} &= -\alpha tr\{(x^T Bx - D)^{s-1}\{x^T Axx^T Bx \\ &\quad - x^T Bxx^T Ax - \alpha x^T Bx(x^T Bx - D)^m\} \\ &= -\alpha tr\{(x^T Bx - D)^{s-1} x^T x(x^T Bx - D)^m\} \leq 0\end{aligned}$$

provided that B is positive definite and $s - 1 + m$ is even.

5 Simulations

Analytical solutions of most of the proposed dynamical systems are not available. Thus simulation offers a way to gain insight in the behavior of these systems. In this simulation, a matrix $A = U\Sigma U^T$ of size $n = 6$ is generated so that Σ is a diagonal matrix with eigenvalues $1, 2, 3, 4, 5, 6$. The matrix U is orthogonal and is generated randomly using the Matlab function *qr* applied to a random matrix. The MCA system (11), with the + sign, is used to compute the 4-dimentional minor components of the matrix A.

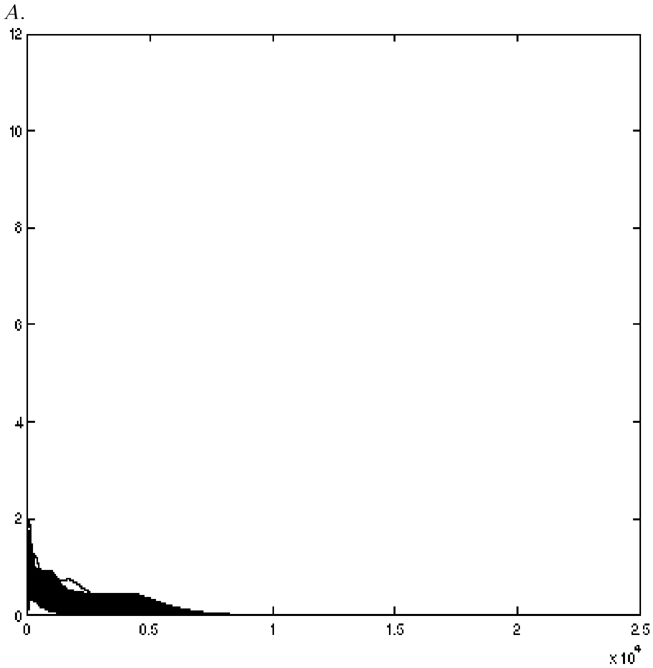

Figure 1: *This plot shows the error $e(k) = ||off(x^T Ax)||_2 + ||off(x^T x)||_2$ for 22000 iterations repeated 100 times using different initial conditions. Note that $e(k)$ eventually diminished indicating that both $x^T Ax$ and $x^T x$ are nearly diagonal.*

The initial condition x_0 is chosen randomly using the Matlab function *rand*, $\alpha = 20$, and the matrix $D = diag\{0.8744, 0.0150, 0.7680, 0.9708\}$. The plots shown in the figure are obtaind by solving (11) 100 times each with 22000 iterations and using the same stepsize $\epsilon = 0.0048$ and the same diagonal matrix D. Convergence is measured by the euclidean norm of the off-diagonal elements of $x^T Ax$ and $x^T x$, i.e., $e(k) = ||off(x^T Ax)||_2 + ||off(x^T x)||_2$, where $||C||_2$ denotes the euclidean norm of C and $off(C)$ denotes the matrix $off(C) = C - diag(C)$. In another similar simulation, it is shown (using the Matlab long fromat) that the smallest four eigenvalues of A which are approximated by the eigenvalues of $x^T Ax(x^T x)^{-1}$ are$\{1.00000000000000, 2.00000000000000, 3.00000000000000, 4.00148086934653\}$.

6 Conclusions

We have proposed a number of dual-purpose learning rules for extracting principle and minor components of general matrices. In these rules, switching a learning rule from being a PCA ordinary differential equation (ODE) to being an MCA ODE is achieved by merely multiplying few terms of the learning rule by -1. It was shown that the proposed ODEs include some existing dual-purpose rules as special cases. Additionally, we have shown that a penalty term can be added so that the resulting systems are globally asymptotically stable. Many simulation results which are not reported here due to space limitation have indicated very good agreement with the theoretical results. A comprehensive analysis of many learning rules stated in Equations (18a)-(18d), (19a)-(19e), and (20a)-(20e), will be detailed in a forthcoming paper.

References

[1] E. Oja, "Principal components, minor components, and linear neural networks. Neural Networks," 5:927-935, November 1992.

[2] T. D. Sanger, "Optimal unsupervised learning in a single-layer linear feedforward network," Neural Networks, 2:459-473, 1989.

[3] L. Xu, "Least mean square error recognition principle for self organizing neural nets," Neural Networks, 6:627-648, 1993.

[4] T. Chen and Shun-ichi Amari, "Unified Stabilization Approach to Principal and Minor Components Algorithms," Neural Networks, Vol. 14, pp. 1377-1387, 2001.

[5] G. Cirrincione, M. Cirrincione, J. Herault, and S. Van Huffel, "The MCA EXIN Neuron for the Minor Component Analysis," IEEE Trans. on Neural Networks, Vol. 13, No. 1, pp. 160-187, January 2002.

[6] J. H. Manton, Uwe Helmke, and Iven M. Y. Mareels, "Dynamical Systems for Principal and Minor Component Analysis," Proceedings of the 42nd IEEE Conference on Decision and Control, Maui, Hawaii USA, pp. 1863-1868, December 2003.

[7] J. LaSalle, "Some Extensions of Liapunov's Second Method," IRE Transactions on Circuit Theory, Volume 7, Issue 4, Dec 1960 Page(s):520-527

[8] M.A. Hasan, "Natural Gradient for Minor Component Extraction," Proceedings of the IEEE International Symposium on Circuits and Systems, 2005 (ISCAS 2005), 23-26 May 2005 Page(s):5138- 5141.

[9] M.A. Hasan, "Diagonally weighted and shifted criteria for minor and principal component extraction," Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005. IJCNN'05, Volume 2, 31 July-4 Aug. 2005 Page(s):1251-1256.

[10] R. Moller, "A self-stabilizing learning rule for minor component analysis," Int. J. Neural Syst., vol. 14, pp. 1-8, 2004.