Continuous-time flows on quotient spaces for principal component analysis

P.-A. Absil*

Abstract

A novel matrix flow for principal or minor component analysis is constructed. The development is based on decompositions of the set of full-rank $n \times p$ matrices into orbits of Lie group actions.

Key words. Principal component analysis, principal subspace analysis, minor component analysis, continuous-time matrix flows, quotient spaces, Lie group actions, semidefinite Lyapunov functions, leftmost and rightmost eigenspaces.

1 Introduction

Given a symmetric matrix A, a flow on $\mathbb{R}^{n\times p}$ is said to have principal subspace analysis (PSA) properties if the column space of the solution converges to the p-dimensional eigenspace of A associated with the largest eigenvalues. If moreover the columns of the solution converge to the p principal eigenvectors of A, i.e. those corresponding to the largest eigenvalues, then the flow is said to achieve principal component analysis (PCA).

Several continuous-time dynamical systems on matrix spaces (also called matrix flows) have been proposed in the literature that achieve PSA or even PCA. Interest for studying continuous-time flows stems in part from the work of Ljung [Lju77] and Kushner and Clark [KC78] relating the behaviour of learning algorithms to the one of associated differential equations; see e.g. Oja and Karhunen [OK85] for an application.

Early analyses of these matrix flows have focused on local stability issues without addressing the problem of global convergence in a mathematically satisfactory way. A breakthrough came with the analysis by Yoshizawa *et al.* [YHS01] of the flow

$$\dot{Y} = AYN - YNY^TAY,\ Y \in \mathbb{R}^{n\times p}. \tag{1}$$

*School of Computational Science and Information Technology, Florida State University, Tallahassee FL 32306-4120, USA (www.csit.fsu.edu/∼absil). This work was supported partially by the National Science Foundation of the USA under Grant ACI0324944 and by the School of Computational Science and Information Technology of Florida State University. Part of this work was done while the author was a Research Fellow with the Belgian National Fund for Scientific Research ("Aspirant du F.N.R.S.") at the University of Liège.

This flow was studied by Brockett [Bro91] in the case where Y belongs to the set of orthonormal $n \times n$ matrices; for the choice $N = I$ it yields the well-known Oja flow [Oja89]. Assuming that A is positive definite, Yoshizawa *et al.* [YHS01] show that (1) is a gradient flow for a certain cost function on $\mathbb{R}^{n\times p}$ endowed with a well-chosen Riemannian metric. Using Łojasiewicz's theorem [Loj84] (or see [KMP00, AMA04]), they show that all solutions of (1) converge to a single equilibrium point. More recently, Manton *et al.* [MHM03a, MHM03b] have proposed a gradient flow that can serve for principal and minor component analysis. A difficulty with gradient-based approaches, however, is the absence of a systematic procedure to detect whether a flow can be expressed as a gradient flow and to determine the corresponding cost function and metric.

In this paper, we propose a constructive procedure that yields a matrix flow with PSA and PCA properties. The technique is to consider two quotients of $\mathbb{R}^{n\times p}$, whose fibers include all matrices with the same column space and all matrices related by rotations, respectively. In the first quotient (the "Grassmann manifold"), the objective is to design a flow that converges to the fiber corresponding to the dominant p-dimensional eigenspace of A; this guarantees PSA properties. In the second quotient (the "shape manifold"), the objective is to converge to the fiber of orthonormal matrices; this ensures that Y stays well-conditioned. As we will show, both objectives can be achieved within the same matrix flow. Moreover, the remaining degrees of freedom, which correspond to rigid rotations of Y within its column space, can be controlled in such a way that the columns of Y converge to the dominant eigenvectors of A; this guarantees PCA properties. The convergence analysis of the flow is based on semidefinite Lyapunov functions and on the invariance properties of ω-limit sets. Interestingly, like in [MHM03a] but in contrast to [YHS01], the matrix A need not be positive definite; the flow converges to the eigenspace of A corresponding the algebraically largest eigenvalues. As a consequence, the flow can also be applied to $-A$ for minor component analysis.

The paper is organized as follows. The intertwined structure of quotients is introduced in Section 2. The Grassmann quotient and the shape quotient are considered separately in Sections 3 and 4. The remaining degrees of freedom are dealt with in Section 5 and the global behaviour of the flow is analysed. The results are summarized in Theorem 5.4.

2 A geometric structure for matrix flows

In this section we show how a matrix differential equation can be decomposed into three terms with a specific geometric interpretation. We will then take advantage of this structure to derive a matrix flow with PSA and PCA properties.

The following definitions and conventions will be used throughout the paper.

Definition 2.1 *Let A be a symmetric $n \times n$ matrix. Let $\lambda_1 \geq \ldots \geq \lambda_n$ be the eigenvalues of A and let $(v_1, \ldots, v_n)$ be a corresponding orthonormal basis of eigenvectors. Let $V = [v_1|\ldots|v_p]$. Then* span(V)*, the column space of V, is a p-dimensional invariant subspace of A corresponding to the eigenvalues of A with largest algebraic value. We call* span(V) *a* rightmost *p-dimensional eigenspace of A. The matrix A has a unique p-dimensional rightmost eigenspace if and only if $\lambda_p > \lambda_{p+1}$. Similarly, we define* leftmost *eigenspaces as those with algebraically smallest eigenvalues.*

Let $\mathbb{R}_*^{n\times p}$ denote the set of full-rank $n \times p$ matrices, and let F be a function on $\mathbb{R}_*^{n\times p}$ into $\mathbb{R}^{n\times p}$. Consider the matrix differential equation

$$\dot{Y}(t) = F(Y(t)). \tag{2}$$

We will assume throughout this paper that for all initial conditions $Y(0) \in \mathbb{R}_*^{n\times p}$, the differential equation (2) has one and only one solution curve $Y(t)$ in $\mathbb{R}_*^{n\times p}$, $t \in [0, +\infty)$.

The ultimate goal of PCA is to choose F such that $Y(t)$ converges (up to sign changes) to the matrix V defined in Definition 2.1. A key point for this paper is to notice that this goal is reached if the following three conditions hold together: (i) the column space of Y converges to the rightmost eigenspace of A; (ii) Y converges to the set of orthonormal matrices; (iii) Y^TAY converges to the set of diagonal matrices with nonincreasing entries.

With the intention of considering these three properties separately, we decompose (2) as follows:

$$\dot{Y} = F(Y) \equiv W_Y + Y(Y^TY)^{-1}S_Y + Y(Y^TY)^{-1}\Omega_Y \tag{3}$$

where W_Y is orthogonal to Y (i.e. $Y^TW_Y = 0$), S_Y is symmetric ($S_Y^T = S_Y$) and Ω_Y is skew-symmetric ($\Omega_Y^T = -\Omega_Y$). The purpose of the subscripts "Y" is to denote the dependency on Y; when possible, they will be omitted for simplicity. The decomposition (3) always exists and is unique; it is easily checked that $W = \Pi_{Y_\perp}F(Y)$, $S = \mathrm{sym}(Y^TF(Y))$ and $\Omega = \mathrm{skew}(Y^TF(Y))$, where

$$\Pi_{Y_\perp} = I - Y(Y^TY)^{-1}Y^T \tag{4}$$

is the orthogonal projector onto the orthogonal complement of the column space of Y, and $\mathrm{sym}(A) = (A + A^T)/2$ and $\mathrm{skew}(A) = (A - A^T)/2$ denote the symmetric and the skew-symmetric parts in the decomposition of a matrix.

The three terms in (3) produce specific actions on the "span" and the "shape" of Y. The *span*, or column space, of Y is the linear subspace spanned by the columns of Y, namely $\mathrm{span}(Y) := \{Ym : m \in \mathbb{R}^p\}$; two matrices Y_1 and Y_2 of $\mathbb{R}_*^{n\times p}$ have the same span if and only if there exists an (invertible) matrix M such that $Y_2 = Y_1M$. We say that two matrices Y_1 and Y_2 have the same *shape* if they are related by a rotation, namely, there exists an orthogonal matrix U such that $Y_2 = UY_1$. It is possible to show that the terms $Y(Y^TY)^{-1}S_Y$ and $Y(Y^TY)^{-1}\Omega_Y$ in (3) do not alter the span of Y, and that the terms W_Y and $Y(Y^TY)^{-1}\Omega_Y$ do not modify the shape of Y. Therefore, the term W in (3) controls the evolution of the span of Y, the term $Y(Y^TY)^{-1}S$ determines the evolution of the shape of Y, and the term $Y(Y^TY)^{-1}\Omega$ controls the remaining degrees of freedom which correspond to rotations that do not modify the span. See illustration on Figure 1.

The principle of our design is as follows: (i) Choose W such that the span of Y converges to the rightmost eigenspace of A, for almost all initial condition in $\mathbb{R}_*^{n\times p}$, independently of S and Ω. (ii) Choose S such that the shape of Y converges to the orthonormal shape, for all initial condition in $\mathbb{R}_*^{n\times p}$, independently of W and Ω.

The natural third condition would be to choose Ω such that Y^TAY converges to the set of diagonal matrices, independently of W and S. Then the combination of the three terms would directly yield a PCA flow. Unfortunately this goal is not achievable: the distribution generated by W and $Y(Y^TY)^{-1}S$ is not integrable and moreover the set $\{Y \in \mathbb{R}_*^{n\times p} : Y^TAY \text{ is diagonal}\}$ has a very complicated structure. Nevertheless, in view of the

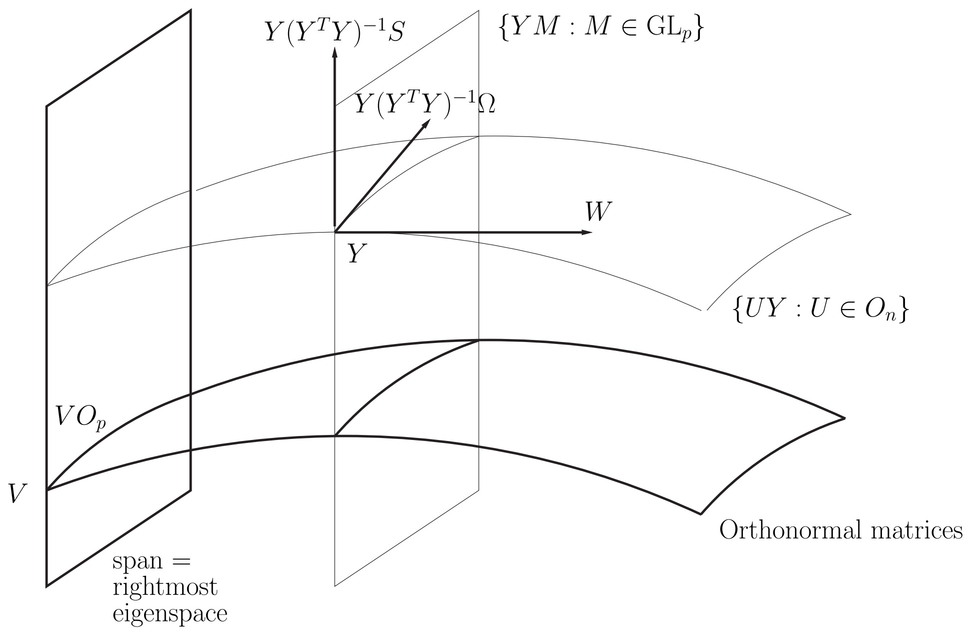

Figure 1: Illustration of the intertwined quotient structures on $\mathbb{R}_*^{n\times p}$.

choices made for W and S, Y approaches the set VO_p of orthonormal matrices that span the rightmost eigenspace of Y. Therefore, we will do the following: (iii) Choose Ω such that for almost all initial condition in the set VO_p of orthonormal matrices spanning the rightmost eigenspace of A, Y converges to V up to sign changes. Convergence results for the flow on all of $\mathbb{R}_*^{n\times p}$ are then obtained using the theory of semidefinite Lyapunov functions and the properties of invariant sets (see Theorem 5.4).

3 Flows on the Grassmann manifold

In this section, we consider task (i) presented in the previous section, i.e. choosing W in (3) so that the span of Y converges to the rightmost eigenspace of A, for almost all initial condition in $\mathbb{R}_*^{n\times p}$, independently of S and Ω. This section and the next ones use basic notions of differential geometry, Lie groups and dynamical systems; the necessary background can be found e.g. in [Boo75], [DK00] and [Kha96].

Let $\mathrm{Grass}(p, n)$, called the *Grassmann manifold*, denote the set of p-dimensional subspaces in $\mathbb{R}^n$. Let $\varphi_t^F Y_0$ denote the unique solution of (3) with the initial condition $Y(0) = Y_0$; φ^F is called the flow of F. We say that (the flow of) F *induces a flow on* $\mathrm{Grass}(p, n)$ if the following condition holds:

$$\mathrm{span}(Y_1) = \mathrm{span}(\hat{Y}_2) \;\Rightarrow \mathrm{span}(\varphi_t^F Y_1) = \mathrm{span}(\varphi_t^{\hat{F}} \hat{Y}_2), \tag{5}$$

for all $t \geq 0$ and all $Y_1, Y_2 \in \mathbb{R}_*^{n\times p}$. More generally, F and another function $\hat{F}$ induce one and the same flow on $\mathrm{Grass}(p, n)$ if

$$\mathrm{span}(Y_0) = \mathrm{span}(\hat{Y}_0) \;\Rightarrow \mathrm{span}(\varphi_t^F Y_0) = \mathrm{span}(\varphi_t^{\hat{F}} \hat{Y}_0), \tag{6}$$

for all $t \geq 0$ and all $Y_0, \hat{Y}_0 \in \mathbb{R}_*^{n\times p}$.

Theorem 3.1 *Let F be decomposed as in* (3)*. Then the flow of F induces a flow on the Grassmann manifold if and only if*

$$W_{YM} = W_Y M \tag{7}$$

for all $Y \in \mathbb{R}_*^{n\times p}$ and all $M \in \mathbb{R}_*^{p\times p}$. Let

$$\hat{F}(Y) \equiv \hat{W}_Y + Y(Y^TY)^{-1}\hat{S}_Y + Y(Y^TY)^{-1}\hat{\Omega}_Y \tag{8}$$

be the same decomposition for $\hat{F}$. Then F and $\hat{F}$ induce one and the same flow on $\mathrm{Grass}(p,n)$ *if and only if* (7) *holds and*

$$W_Y = \hat{W}_Y \tag{9}$$

for all $Y \in \mathbb{R}_*^{n\times p}$.

Note that these conditions involve neither S nor Ω; if F induces a flow on $\mathrm{Grass}(p,n)$, then this flow is independent of S and Ω.

Our proof of Theorem 3.1, see [AMS04a], is based on the fact that $\mathrm{Grass}(p,n)$ is the quotient $\mathbb{R}_*^{n\times p}/\psi_1$ of $\mathbb{R}_*^{n\times p}$ by the Lie group action

$$\psi_1 : \mathrm{GL}_p \times \mathbb{R}_*^{n\times p} \to \mathbb{R}_*^{n\times p} : (M,Y) \mapsto YM$$

where $\mathrm{GL}_p := \mathbb{R}_*^{p\times p}$ is the general linear group. The orbits of the group action ψ_1 comprise all the matrices with the same span (see e.g. [AMS04b] for details).

Now take

$$W_Y := \Pi_{Y_\perp} AY \tag{10}$$

where A is as in Definition 2.1 and $\Pi_{Y_\perp}$ is defined in (4). It is easily checked that W satisfies condition (7). Therefore, in view of Theorem 3.1, the dynamical system

$$\dot{Y} = F(Y) \equiv \Pi_{Y_\perp} AY + Y(Y^TY)^{-1}S_Y + Y(Y^TY)^{-1}\Omega_Y, \tag{11}$$

$S_Y^T = S_Y$, $\Omega_Y^T = -\Omega_Y$, induces a flow on $\mathrm{Grass}(p,n)$ which depends neither on S nor on Ω.

It remains to show that this Grasmmann flow has PSA properties. To this end, consider the Rayleigh quotient cost function $f(Y) = \mathrm{trace}((Y^TY)^{-1}Y^TAY)$. It defines a cost function on $\mathrm{Grass}(p,n)$. The flow on Grassmann induced by (11) is the gradient ascent flow of f with respect to the canonical metric on Grassmann; see [HM94]. Since the Grassmann manifold is compact, it follows from Łojasiewicz's theorem [Loj84] (which admits a generalization on Riemannian manifolds, see Lageman [Lag02]) that the flow on $\mathrm{Grass}(p,n)$ converges to an eigenspace of A. Moreover, the rightmost eigenspace of A is the only one to be stable, as the others are saddle points or minima of f. This rationale leads to the following result.

Lemma 3.2 *Consider the dynamics* (11) *where A is a fixed $n \times n$ symmetric matrix. Then the span of Y converges to an eigenspace of A. Only the rightmost eigenspace of A (see Definition 2.1) is stable.*

4 Flows on the shape manifold

We now consider task (ii) presented in Section 2, namely, choose S in (3) such that the shape of Y converges to the orthonormal shape, for all initial condition in $\mathbb{R}_*^{n\times p}$, independently of W and Ω.

This task is tackled much as task (i). Consider the Lie group action

$$\psi_2 : O_n \times \mathbb{R}_*^{n\times p} \to \mathbb{R}_*^{n\times p} : (U, Y) \mapsto UY$$

where $O_n = \{U \in \mathbb{R}^{n\times n} : U^T U = I_n\}$ is the orthogonal group. The orbits of this group action correspond to all the full-rank $n \times p$ matrices with same shape, i.e. related by rotations. The *shape manifold* is the quotient $\mathbb{R}_*^{n\times p}/\psi_2$. We say that F and $\hat{F}$ *induce one and the same flow on the shape manifold* if the following condition holds:

$$\hat{Y}_0 = U_0 Y_0, U_0 \in O_n \Rightarrow \forall t \geq 0 : \exists U_t \in O_n : \varphi_t^{\hat{F}} \hat{Y}_0 = U_t \varphi_t^F Y_0. \tag{12}$$

Theorem 4.1 *Let F be decomposed as in* (3)*. Then the flow of F induces a flow on the shape manifold if and only if*

$$S_{UY} = S_Y \tag{13}$$

for all $Y \in \mathbb{R}_*^{n\times p}$ and all $U \in O_n$. Let $\hat{F}$ be decomposed as in (8)*. Then F and $\hat{F}$ induce one and the same flow on* $\mathrm{Grass}(p, n)$ *if and only if* (13) *holds and*

$$S_Y = \hat{S}_Y \tag{14}$$

for all $Y \in \mathbb{R}_*^{n\times p}$.

Note that these conditions only involve S, and neither W nor Ω.

Now consider the choice

$$S_Y := I_p - Y^T Y. \tag{15}$$

It satisfies condition (13). Therefore, in view of Theorem 4.1, the dynamical system

$$\dot{Y} = F(Y) \equiv W_Y + Y(Y^T Y)^{-1}(I_p - Y^T Y) + Y(Y^T Y)^{-1}\Omega_Y, \tag{16}$$

$Y^T W_Y = 0$, $\Omega_Y^T = -\Omega_Y$, induces a flow on the shape manifold which does not depend on W and Ω. To see that the flow (16) converges to the set of orthonormal matrices for all initial condition in $\mathbb{R}_*^{n\times p}$, consider the function $v := \|I - Y^T Y\|_F^2 := \mathrm{trace}((I - Y^T Y)(I - Y^T Y))$; one obtains $\dot{v} = -4v$, hence $v(t)$ converges exponentially to zero independently of the initial condition Y. We have thus shown the following:

Lemma 4.2 *Let A be a symmetric $n \times n$ matrix. Let $Y(t)$ be the solution of* (16) *with $Y(0) = Y_0 \in \mathbb{R}_*^{n\times p}$. Then $Y(t)$ approaches the set of orthonormal matrices as $t \to +\infty$.*

5 A new PCA flow

It follows from the two previous sections that the flow

$$\dot{Y} = F(Y) \equiv \Pi_{Y_\perp} AY + Y(Y^TY)^{-1}(I_p - Y^TY) + Y(Y^TY)^{-1}\Omega_Y, \tag{17}$$

$\Omega_Y^T = -\Omega_Y$, satisfies properties (i) and (ii) put forward in Section 2. That is, the solution $Y(t)$ of (17) approaches the set of orthonormal matrices that span the rightmost eigenspace of A (unless $Y(t)$ converges to an unstable set).

The remaining degrees of freedom in Ω correspond to motions along the intersections of the orbits of ψ_1 and ψ_2, which are orbits of

$$\psi_3 : O_p \times \mathbb{R}_*^{n\times p} \to \mathbb{R}_*^{n\times p} : (Q, Y) \mapsto Y(Y^TY)^{-1/2}Q(Y^TY)^{1/2}. \tag{18}$$

Consider the choice

$$\Omega := [Y^TAY, N] \tag{19}$$

where $[B, C] = BC - CB$ and

$$N := \mathrm{diag}(p, \ldots, 1). \tag{20}$$

This yields the dynamical system

$$\dot{Y} = F(Y) \equiv \Pi_{Y_\perp} AY + Y(Y^TY)^{-1}(I_p - Y^TY) + Y(Y^TY)^{-1}[Y^TAY, N]. \tag{21}$$

The convergence results for (21) are given in Theorem 5.4 and are based on a few lemmas.

The result below is a consequence of the behaviour of the double bracket flow; see e.g. [HM94, 2.1].

Lemma 5.1 *Let A, $\lambda_1, \ldots, \lambda_n$, $v_1, \ldots, v_n$ and V be as in Definition 2.1. Assume that $\lambda_1 > \ldots > \lambda_{p+1}$. Consider the dynamics (21) with $N = \mathrm{diag}(p, \ldots, 1)$. Then $VO_p := \{VQ : Q \in O_p\}$ is an invariant set with respect to (21). Consider (21) restricted to VO_p. For all initial condition, $Y(t)$ converges to a matrix W of eigenvectors of A, i.e. $W = VP$ where P is a signed permutation matrix. The equilibrium point W is (asymptotically) stable conditionally to VO_p if and only if $W = V\mathrm{diag}(\pm 1, \ldots, \pm 1)$.*

We now consider the dynamics (21) and show that $V\mathrm{diag}(\pm 1, \ldots, \pm 1)$ is stable on all of $\mathbb{R}_*^{n\times p}$, and not only conditionally to VO_p. The proof is based on semidefinite Lyapunov functions; see [IKO96] or [SJK97, Section 2.3.2].

Lemma 5.2 *Let A, V, $\lambda_1, \ldots, \lambda_n$, $v_1, \ldots, v_n$ and N be as in Lemma 5.1. Then V is asymptotically stable for the dynamics (21) in $\mathbb{R}_*^{n\times p}$.*

Proof. Consider

$$v(Y) = \sum_{i=1}^{p} \lambda_i - \mathrm{trace}((Y^TY)^{-1}Y^TAY) + \mathrm{trace}((I - Y^TY)(I - Y^TY)). \tag{22}$$

One shows that
(i) $v(Y) \geq 0$ for all $Y \in \mathbb{R}_*^{n\times p}$ and $v(V) = 0$.

(ii) $\dot{v}(Y) \leq 0$ for all $Y \in \mathbb{R}_*^{n\times p}$.
(iii) V is asymptotically stable conditionally to $\{Y \in \mathbb{R}_*^{n\times p} : \dot{v}(Y) = 0\}$.
The conclusion comes from the theory of semidefinite Lyapunov functions, in particular [IKO96, corollary 1]. □

Finally, we need a technical result about the limit sets of trajectories of dynamical systems. By definition (see e.g. [Kha96]), the ω-limit set of a trajectory $x(t)$ is the set of points p for which there exists a sequence $\{t_n\}$, with $t_n \to +\infty$ as $n \to +\infty$, such that $x(t_n) \to p$ as $n \to +\infty$.

Lemma 5.3 *Suppose that*

$$\dot{x} = f(x) \tag{23}$$

has one and only one solution $x(t)$ for all $t \geq 0$ with initial condition $x(0) = x_0 \in \mathbb{R}^n$. If $\bar{x}$ is a stable point for (23) and $\bar{x}$ belongs to the ω-limit set $L^+(x_0)$ of $x(t)$, then $L^+(x_0) = \{\bar{x}\}$, i.e. $x(t)$ converges to $\bar{x}$.

Proof. Let $\epsilon > 0$. By the definition of stability, there exists $\delta > 0$ such that if $x(t^*) \in B_\delta(\bar{x})$ then $x(t) \in B_\epsilon(\bar{x})$ for all $t \geq t^*$. Since $\bar{x} \in L^+(x_0)$, there exists t^* such that $x(t^*) \in B_\delta(\bar{x})$. In conclusion, there exists t^* such that $x(t) \in B_\epsilon(\bar{x})$ for all $t \geq t^*$. This means that $x(t)$ converges to $\bar{x}$. □

We are now ready to state the main result of this paper:

Theorem 5.4 *Let A be an $n \times n$ symmetric matrix with eigenvalues $\lambda_1 \geq \ldots \geq \lambda_n$ and associated orthonormal eigenvectors $v_1, \ldots, v_n$. Let $V = [v_1|\ldots|v_p]$. Let $N = \mathrm{diag}(p, \ldots, 1)$. Consider the dynamics*

$$\dot{Y} = \Pi_{Y_\perp} AY + Y(Y^TY)^{-1}(I_p - Y^TY) + Y(Y^TY)^{-1}[Y^TAY, N] \tag{24}$$

with full-rank $n \times p$ initial condition $Y(0) = Y_0 \in \mathbb{R}_*^{n\times p}$ and $\Pi_{Y_\perp}$ defined as in (4). Then
(i) $Y(t)^TY(t) \to I_p$ as $t \to +\infty$.
(ii) The flow (24) induces a subspace flow, i.e. span$(Y(t))$ *only depends on* span$(Y(0))$.
(iii) There exists and eigenspace $\mathcal{S}$ of A such that span$(Y(t)) \to \mathcal{S}$ *as $t \to +\infty$.*
(iv) The eigenspace $\mathcal{S}$ is asymptotically stable for the induced subspace flow if and only if it is the unique p-dimensional rightmost eigenspace of A (see Definition 2.1). $\mathcal{S}$ is unstable if it is not a rightmost eigenspace of A.
Now assume that $\lambda_1 > \ldots > \lambda_{p+1}$.
(v) The set $VO_p = \{VQ : Q^TQ = I_p\}$ is invariant with respect to (24). If the initial condition Y_0 is in VO_p, then $Y(t)$ converges to an orthonormal matrix $\tilde{V}$ whose columns are eigenvectors of A. The equilibrium point $\tilde{V}$ is stable conditionally to VO_p if and only if $\tilde{V} = V\mathrm{diag}(\pm 1, \ldots, \pm 1)$.
(vi) Let $\mathcal{W}_1$ be the set of initial conditions $Y_0 \in VO_p$ such that $Y(t)$ converges to a matrix $V\mathrm{diag}(\pm 1, \ldots, \pm 1)$ and let $\mathcal{W}_2$ be the complement of $\mathcal{W}_1$ in VO_p. If span$(Y(t))$ *converges to the rightmost eigenspace (generic case) then, either $L^+(Y_0) \subseteq \mathcal{W}_2$, or $Y(t)$ converges to a $V\mathrm{diag}(\pm 1, \ldots, \pm 1)$.*

Proof. Only (v) still needs to be proven. Assume that $L^+(Y_0) \not\subseteq \mathcal{W}_2$. Then there exists $Y_* \in \mathcal{W}_1 \cap L^+(Y_0)$. Say $L^+(Y_*) = \{V\}$ without loss of generality. Since $L^+(Y_0)$ is invariant and $L^+(Y_0) \ni Y_*$ and $L^+(Y_*) = \{V\}$, it follows that $V \in L^+(Y_0)$. In view of Lemma 5.2, V is stable for the dynamics (24). It then follows from Lemma 5.3 that $Y(t) \to V$. □

6 Conclusion

Using an approach based on quotient spaces and semidefinite Lyapunov functions, we have constructed the matrix flow (24) with principal and minor component analysis properties. Detailed convergence results have been presented in Theorem 5.4.

Acknowledgements

The author would like to thank R. Sepulchre, R. Mahony, J. H. Manton, G.-B. Stan and Ch. Lageman for useful discussions related to the topic of this paper. Special thanks to K. Gallivan and J. Jordan for their suggestions that considerably improved the quality of this paper.

References

[AMA04] P.-A. Absil, R. Mahony, and B. Andrews, *Convergence of the iterates of descent methods for analytic cost functions*, submitted to SIAM J. Optim., March 2004.

[AMS04a] P.-A. Absil, R. Mahony, and R. Sepulchre, *Continuous-time subspace flows related to the symmetric eigenproblem*, submitted to Linear Algebra Appl., April 2004.

[AMS04b] ———, *Riemannian geometry of Grassmann manifolds with a view on algorithmic computation*, Acta Appl. Math. **80** (2004), no. 2, 199–220.

[Boo75] W. M. Boothby, *An introduction to differentiable manifolds and Riemannian geometry*, Academic Press, 1975.

[Bro91] R. W. Brockett, *Dynamical systems that sort lists, diagonalize matrices, and solve linear programming problems*, Linear algebra appl. **146** (1991), 79–91.

[DK00] J. J. Duistermaat and J. A. Kolk, *Lie groups*, Springer-Verlag, Berlin, 2000.

[HM94] U. Helmke and J. B. Moore, *Optimization and dynamical systems*, Springer, 1994.

[IKO96] A. Iggidr, B. Kalitine, and R. Outbib, *Semidefinite Lyapunov functions: stability and stabilization*, Math. Control Signals Systems **9** (1996), no. 2, 95–106.

[KC78] H. J. Kushner and D. S. Clark, *Stochastic approximation methods for constrained and unconstrained systems*, Applied Mathematical Sciences, vol. 26, Springer-Verlag, 1978.

[Kha96] H. K. Khalil, *Nonlinear systems, second edition*, Prentice Hall, 1996.

[KMP00] K. Kurdyka, T. Mostowski, and A. Parusiński, *Proof of the gradient conjecture of R. Thom*, Ann. of Math. (2) **152** (2000), no. 3, 763–792.

[Lag02] C. Lageman, *Konvergenz reell-analytischer gradientenähnlicher Systeme*, Diplomarbeit im Fach Mathematik, Mathematisches Institut, Universität Würzburg (Betreuer: Prof. Dr. Uwe Helmke), Januar 2002.

[Lju77] L. Ljung, *Analysis of recursive stochastic algorithms*, IEEE Trans. Automatic Control **22** (1977), no. 4, 551–575.

[Loj84] S. Lojasiewicz, *Sur les trajectoires du gradient d'une fonction analytique*, Seminari di Geometria 1982-1983 (Università di Bologna, Istituto di Geometria, Dipartimento di Matematica), 1984, pp. 115–117.

[MHM03a] J. H. Manton, U. Helmke, and I. M. Y. Mareels, *A dual purpose principal and minor component flow*, submitted to Systems and Control Letters, 2003.

[MHM03b] ______, *Dynamical systems for principal and minor component analysis*, Proceedings of the IEEE International Conference on Decision and Control, 2003.

[Oja89] E. Oja, *Neural networks, principal components, and subspaces*, Int. J. Neural Syst. **1** (1989), 61–68.

[OK85] E. Oja and J. Karhunen, *On stochastic approximation of the eigenvectors and eigenvalues of the expectation of a random matrix*, J. Math. Anal. Appl. **106** (1985), no. 1, 69–84.

[SJK97] R. Sepulchre, M. Jankovic, and P. V. Kokotović, *Constructive nonlinear control*, Springer-Verlag, 1997.

[YHS01] S. Yoshizawa, U. Helmke, and K. Starkov, *Convergence analysis for principal component flows*, Int. J. Appl. Math. Comput. Sci. **11** (2001), no. 1, 223–236.