WA05 11:40

Proceedings of the 34th
Conference on Decision & Control
New Orleans, LA - December 1995

DYNAMIC INVERSION AND POLAR DECOMPOSITION OF MATRICES

Neil H. Getz* and Jerrold E. Marsden†

ABSTRACT

Using the recently introduced concept of a "dynamic inverse" of a map, along with its associated analog computational paradigm, we construct continuous-time nonlinear dynamical systems which produce both regular and generalized inverses of time-varying and fixed matrices, as well as polar decompositions.

1. Introduction

In [1] (see this proceedings) we introduced a technique in which a dynamical system is used to generate an approximation to the solution $\theta_*(t)$ of a nonlinear vector equation of the form $F(\theta, t) = 0$. As we saw in Example 4.1 of [1], one may also pose the inverse of a time-varying matrix as a solution to an equation of the form $F(\Gamma, t) = 0$. Square roots and other matrix functions may be posed similarly. Motivated by this realization, in this paper[1] we will further investigate the use of dynamic inversion to construct dynamical systems that perform matrix inversion as well as polar decomposition.

Dynamical methods of matrix inversion have appeared in the neural network literature [4, 5]. We will show in Section 3.1 that these neural network methods may be regarded as a special case of dynamic inversion. A decomposition related to polar decomposition has also appeared in Helmke and Moore [6], though, as the authors point out, their method does not guarantee the positive definiteness of the symmetric component of the polar decomposition. The approaches of [4, 5, 6] are gradient methods and produce exact results asymptotically. In this paper, using dynamic inversion we will derive a system that produces the desired inverse and polar decomposition products at any preassigned time $t_1 > 0$.

In Example 4.1 of [1] we examined the application of dynamic inversion to the problem of inverting time-varying matrices where we assumed that a good approximation existed for the inverse of the time-varying matrix at an initial time. In Section 2 we will show some further applications of time-varying matrix inversion. Motivated by the desire to obtain such initial inverses dynamically, in Section 3 we will consider the problem of inverting fixed matrices. By using a matrix homotopy from the identity we will recruit the results of Section 2 to produce exact inversion of positive definite fixed matrices in finite time. In Section 4 we will construct a dynamic inverter which produces the polar decomposition of a time-varying matrix. In Section 5 we revisit the problem of fixed matrix inversion and show how, combining homotopy with dynamic polar decomposition, we may dynamically produce the polar decomposition products as well as the inverse of *any* fixed matrix in finite time without requiring an initial guess at the inverse.

* Electrical Engineering and Computer Sciences, University of California at Berkeley, Berkeley, California 94720, getz@eecs.berkeley.edu.

† Control and Dynamical Systems, California Institute of Technology, Pasadena, California, marsden@cds.caltech.edu.

[1] This paper is a condensed version of [2]. See also [3], Chapter 3.

2. Inverting Time-Varying Matrices

We summarize the results of Example 4.1 of [1] in the following theorem.

Theorem 2.1 *Let $A(t) \in GL(n, \mathbb{R})$, where $GL(n, \mathbb{R})$ is the group of $n \times n$ real nonsingular matrices, be C^1 in t, with $A(t)$, $A(t)^{-1}$, and $\dot{A}(t)$ uniformly bounded in t for all $t \geq 0$. There exists an $r > 0$ such that the following holds: Let $G[w, \Gamma, t]$ be a dynamic inverse (see Definition 2.1 of [1]) of $F(\Gamma, t) = A(t)\Gamma - I$ for all Γ such that $\Gamma - \Gamma_*$ is in $\mathcal{B}_r$, and for all $t \in \mathbb{R}_+$. Let $\Gamma(t) \in \mathbb{R}^{n \times n}$ be the solution to*

$$\dot{\Gamma} = -\mu G[A(t)\Gamma - I, \Gamma, t] - \Gamma \dot{A}(t) \Gamma \tag{1}$$

with $\|\Gamma(0) - \Gamma_*(0)\| \leq r < \infty$. There exists a $\tilde{\mu} > 0$ and a $k > 0$ such that for all $\mu > \tilde{\mu}$, $\|\Gamma(t) - \Gamma_*(t)\|_2 \leq re^{-kt}$ for all $t \geq 0$. In particular $\lim_{t\to\infty} \Gamma = A(t)^{-1}$. ◊

Example 2.2 Consider an n-dimensional mechanical system modeled by the implicit second order differential equation

$$M(q)\ddot{q} + N(q, \dot{q}) = 0. \tag{2}$$

Assume that the matrix $M(q)$ is positive definite and symmetric for all q. It is often convenient, to express such systems in an explicit form, with $\ddot{q}$ alone on the

0-7803-2685-7/95 $4.00 © 1995 IEEE

left side of a second order ordinary differential equation. We will invert $M(q)$ dynamically.

Let $\Gamma = \Gamma^T \in \mathbb{R}^{n\times n}$ be an estimator for M^{-1}. Suppose we know $M^{-1}(q(0))$ approximately. If our approximation is sufficiently close to the true value of $M^{-1}(q(0))$, then setting $\Gamma(0)$ to that approximation, and letting $\mu > 0$ be sufficiently large allows us to apply Theorem 2.1. Then system

$$\begin{aligned}\dot{\Gamma} &= -\mu\Gamma\left(M(q)\Gamma - I\right) - \Gamma\left[\frac{\partial M_{i,j}(q)}{\partial q}\dot{q}\right]_{i,j\in\underline{n}} \cdot \Gamma \\ \ddot{q} &= \Gamma N(q,\dot{q})\end{aligned} \tag{3}$$

provides an exponentially convergent estimate of $\ddot{q}$ for all t. △

3. Inversion of Fixed Matrices

In this section we consider two methods for the dynamic inversion of fixed matrices. In Section 5, relying on the methods of Section 4, we will consider another approach to the same problem.

Fixed matrices may be inverted in a manner similar to the inversion of time-varying matrices as described in the last section. Let Γ denote the estimator for the inverse of a fixed matrix M, with $\Gamma_* = M^{-1}$ the solution for $F(\Gamma) = 0$ where

$$F(\Gamma) := M\Gamma - I. \tag{4}$$

In the case of fixed matrix inversion the estimator for $\dot{\Gamma}_*$ is zero. As a consequence, if Γ is sufficiently close to Γ_*, then we can let $G_1(w,\Gamma) := \Gamma \cdot w$ and use the dynamic inverter

$$\dot{\Gamma} = -\mu\Gamma(M\Gamma - I). \tag{5}$$

Again, we must choose $\Gamma(0)$ close enough to $\Gamma_*(0)$ because $G[w,\Gamma]$ fails to be a dynamic inverse when Γ is singular. Choosing $\Gamma(0)$ sufficiently close to $\Gamma_*(0)$ assures us that, as Γ flows to $\Gamma_* = M^{-1}$, Γ will not pass through the set of singular matrices.

3.1 A Comment on Gradient Methods

As shown in the last section, the dynamic inverse $\Gamma\cdot w$ is not our only choice of a dynamic inverse $G(w,\Gamma,t)$ which is linear in w. It is easily verified that $G[w] = M^T \cdot w$, $w \in \mathbb{R}^{n\times n}$, is also a dynamic inverse for $F(\Gamma) := M\Gamma - I$, and that for this choice of dynamic inverse we do not need to worry about the dynamic inverse becoming singular. It is valid globally and leads to the dynamic inverter

$$\dot{\Gamma} = -\mu M^T(M\Gamma - I). \tag{6}$$

If M is injective, with $M \in \mathbb{R}^{m\times n}$, $m \geq n$, then the equilibrium solution Γ_* of (6) is the left inverse $(M^TM)^{-1}M^T$ of M. If instead we were to choose $F(\Gamma) := \Gamma M - I$ and $G[w] := w \cdot M^T$, and if $M \in \mathbb{R}^{m\times n}$, $m \leq n$, is surjective, then the solution Γ_* would be the right inverse $M^T(MM^T)^{-1}$ of M. The dynamic inverter (6) is the standard least squares gradient flow (see [6], Section 1.6) for the function $\Phi : \mathbb{R}^n \to \mathbb{R}$; $\Gamma \mapsto \Phi(\Gamma)$ where

$$\Phi(\Gamma) := \frac{1}{2}\|M\Gamma - I\|_2^2. \tag{7}$$

It is also the neural-network fixed matrix inverter of Wang [5]. Of course other gradient schemes may have the same solution though they may start from gradients of functions other than $\frac{1}{2}\|M\Gamma - I\|_2^2$ (See, for instance [4]). In general, artificial neural networks are constructed so as to dynamically solve for the minimum of an energy function having a unique (at least locally) minimum, i.e. they realize gradient flows.

3.2 Dynamic Matrix Inversion in Finite Time

The dynamic matrix inverters (5) and (6) above have the potential disadvantage of producing an exact inverse only asymptotically as $t \to \infty$. To correct this we now consider another method. If we could create a time-varying matrix $H(t)$ that is invertible by inspection at $t = 0$, and that equals M at some known $t > 0$, say $t = 1$, then perhaps we could use the technique of Section 2 to invert $H(t)$. Then the solution of the dynamic inverter at time $t = 1$ would be M^{-1}. We require, of course, that $H(t)$ remain in $GL(n,\mathbb{R})$ as t goes from 0 to 1. One ideal candidate for the initial value of the time varying matrix is the identity matrix I, since it is its own inverse.

Example 3.1 Let M be a fixed matrix in $\mathbb{R}^{n\times n}$. We wish to dynamically determine the inverse of M. Let

$$H(t) = h(t, I, M) = (1-t)I + tM. \tag{8}$$

In the space of $n \times n$ matrices, $t \mapsto H(t)$ describes a t-parameterized line segment of matrices from the identity to $M = H(1)$. From the last section we know how to dynamically invert a time-varying matrix given that we have an approximation of its inverse at time $t = 0$. In the present case the inverse at time $t = 0$ is just the identity I. We may invert $H(t)$ by letting $G[w,\Gamma,t] := \Gamma \cdot w$ and substituting $H(t)$ for $A(t)$, and $\dot{H}(t) = M - I$ for $\dot{A}(t)$ in (1). This gives

$$\dot{\Gamma} = -\mu\Gamma\left(H(t)\Gamma - I\right) - \Gamma(M - I)\Gamma. \tag{9}$$

If $\Gamma(0) = I$, then $\Gamma(1) = M^{-1}$. *That is, of course, if $H(t)$ remains nonsingular as t goes from 0 to 1!* △

Remark 3.2 The scheme of Example 3.1 requires that there is no $\lambda \in [0,1]$ for which $h(\lambda, I, M)$ is singular. In other words, the curve $t \mapsto H(t)$ must never leave a connected open subset of $GL(n,\mathbb{R})$. Recall that there are two connected open subsets which comprise $GL(n, \mathbb{R})$; $GL^+(n) = \{M \in \mathbb{R}^{n\times n} | \det(M) > 0\}$ and $GL^-(n) = \{M \in \mathbb{R}^{n\times n} | \det(M) < 0\}$. These two sets are disjoint and are separated by the codimension-1 manifold of singular $n \times n$ matrices. $\triangle$

In the following lemma we give specifies sufficient conditions on M for $h(\lambda, I, M)$ to avoid singularity as λ goes from 0 to 1.

Lemma 3.3 *If $M \in GL(n, \mathbb{R})$ has no eigenvalues in $(-\infty, 0)$, then for each $\lambda \in [0, 1]$, $h(\lambda, I, M)$ is in $GL(n, \mathbb{R})$.* $\diamond$

Note. If M is a positive definite symmetric matrix, then the assumptions of Lemma 3.3 hold.

Proof: Suppose that $h(\lambda, I, M) \equiv (1 - \lambda)I + \lambda M$ is singular for some $\bar{\lambda} \in [0, 1]$. The identity I is nonsingular as is M by assumption, so $\bar{\lambda} \notin \{0, 1\}$. Thus there exists a non-zero $v \in \mathbb{R}^n$ such that

$$((1 - \bar{\lambda})I + \bar{\lambda}M)\, v = 0. \tag{10}$$

Since $\bar{\lambda} \neq 0$ we can divide (10) by $-\bar{\lambda}$ to obtain

$$\left(\frac{(\bar{\lambda} - 1)}{\bar{\lambda}} I - M\right) v = 0. \tag{11}$$

But $\bar{\lambda}$ can only satisfy (11) if $\nu(\bar{\lambda}) := (\bar{\lambda} - 1)/\bar{\lambda}$ is an eigenvalue of M. As λ ranges over $(0, 1)$, $\nu(\lambda)$ ranges over $(-\infty, 0)$. But by assumption M has no eigenvalues in $(-\infty, 0]$. Hence no such λ exists in $(0, 1)$ and $h(\lambda, I, M)$ is nonsingular on $[0, 1]$. $\square$

We may obtain the *exact* inverse of M at any fixed time $t_1 > 0$ by a slight modification of the homotopy. We summarize our results of this section in the following theorem.

Theorem 3.4 *For any fixed positive definite $M \in GL(n, \mathbb{R})$, and any $t_1 > 0$, the solution $\Gamma(t)$ of the dynamic inverter*

$$\dot{\Gamma} = -\mu\Gamma\left(\left((1 - \frac{t}{t_1})I + \frac{t}{t_1}M\right)\Gamma - I\right) - \Gamma(M - I)\Gamma \tag{12}$$

with $\Gamma(0) = I$, satisfies $\Gamma(t_1) = M^{-1}$. $\diamond$

Example 3.5 Let $L \in \mathbb{R}^{m\times n}$ with $m \leq n$ be surjective. The right inverse of L in the Euclidean metric on $\mathbb{R}^n$ is given by $x = L^T(LL^T)^{-1}$. To obtain the right inverse we may apply Theorem 3.4 replacing M by LL^T. Then $L^T(LL^T)^{-1} = L^T\Gamma(1)$. When L is *injective,* the left inverse $(L^TL)^{-1}L^T$ may be obtained by substituting L^TL for M in Theorem 3.4. $\triangle$

By appealing to the polar decomposition in Section 5 below, we will show that we may, at the cost of a slight increase in complexity, use dynamic inversion to produce an exact inverse of *any* invertible M, irrespective of its spectrum, by any fixed time $t_1 \geq 0$.

4. Polar Decomposition for TimeVarying Matrices

In this section we will show how dynamic inversion may be used to perform polar decomposition and inversion of a time-varying matrix. We will assume that $A(t) \in GL(n, \mathbb{R})$, and that $A(t)$, $\dot{A}(t)$, and $A(t)^{-1}$ are bounded for $t \in [0, \infty)$.

Consider the **polar decomposition** [7] of a *fixed* matrix $M \in GL(n, \mathbb{R})$, $M = PU$ where U is in the space of $n \times n$ real orthogonal matrices $O(n, \mathbb{R})$, and P is the symmetric positive definite square root of MM^T. Regarding M as a linear operator $\mathbb{R}^n \to \mathbb{R}^n$, the polar decomposition expresses the action of M on a vector as a rotation (possibly with a reflection) followed by a scaling along the eigenvectors of MM^T. If $M \in GL(n, \mathbb{R})$, then P and U are unique. Matrices of the form MM^T, where M is nonsingular, have only positive real eigenvalues. Thus $A(t)A(t)^T$ is positive definite. For any t, the unique positive definite solution Γ_* to $\Gamma A(t)A(t)^T\Gamma - I = 0$ is $P(t)^{-1}$. Now having $P(t)^{-1}$, from $A(t) = P(t)u(t)$ we get $U(t) = P(t)^{-1}A(t)$, $P(t)^2 = A(t)A(t)^T$, and $P(t) = P(t)^{-1}A(t)A(t)^T$.

Since $P(t)$ is a symmetric $n \times n$ matrix, it is parameterized by $s(n) := n(n + 1)/2$ elements. We will construct the dynamic inverter that produces $P^{-1}(t)$, which is also positive definite and symmetric, in the space $\mathbb{R}^{s(n)}$. Choose an ordered basis $\beta = \{\beta_i\}_{i\in s(n)}$ for the $n \times n$ real-valued symmetric matrices $S(n, \mathbb{R})$. For any $x \in \mathbb{R}^{s(n)}$ there corresponds a unique matrix $x^{\mathrm{m}} \in S(n, \mathbb{R})$ where the correspondence is through the expansion of x^{m} in the ordered basis β,

$$x^{\mathrm{m}} = \sum_{i\in s(n)} x_i\beta_i \in S(n, \mathbb{R}). \tag{13}$$

Conversely, for any $X \in S(n, \mathbb{R})$, let X^{v} denote the vector of the expansion coefficients of

$$X = \sum_{i\in s(n)} x_i\beta_i \tag{14}$$

in the basis β so that $X^{\mathrm{v}} = x$. Then

$$(X^{\mathrm{v}})^{\mathrm{m}} = x^{\mathrm{m}} = X. \tag{15}$$

Let $F : \mathbb{R}^{s(n)} \times \mathbb{R}_+ \to \mathbb{R}^{s(n)}$; $(x, t) \mapsto F(x, t)$ be defined by

$$F(x, t) := (x^{\mathrm{m}}\Lambda(t)x^{\mathrm{m}} - I)^{\mathrm{v}}. \tag{16}$$

where $\Lambda(t) := A(t)A(t)^T$. Let x_* be a solution of $F(x, t) = 0$. Then x_*^{m} is a symmetric square root of $\Lambda(t)$.

Nothing in the form of $F(x, t)$ enforces the *positive definiteness* of the solution $x_*^{\mathrm{m}}(t)$. For instance, for each solution $x_*^{\mathrm{m}}(t)$, $-x_*^{\mathrm{m}}(t)$ is also a solution. Each solution $t \mapsto x_*(t)$ is, however, isolated as long as $D_1F(x_*,t)$ is nonsingular. We will show below that the nonsingularity of $A(t)$ implies the nonsingularity of $D_1F(x_*,t)$. Thus if $x(0)$ is sufficiently close to $(P(0)^{-1})^{\rm v}$, then $x(t) \to (P(0)^{-1})^{\rm v}$ exponentially. Therefore, we can and will enforce the positive definiteness of $x(t)^{\rm m}$ by choice of initial conditions.

4.1 The Lyapunov Map

We will use a linear dynamic inverse for $F(x,t)$ based upon the matrix inverse of $D_1F(x_*,t)$. We will estimate this matrix inverse using dynamic inversion. It is not immediately obvious, however, that $D_1F(x_*,t)$ is invertible. In this subsection we will consider the invertibility of $D_1F(x_*,t)$.

Let $F^l(X,t) := F(X^{\rm v},t)$. Then Differentiating

$$F^l(X,t) = X\Lambda(t)X - I \tag{17}$$

with respect to X gives the map

$$L_{\Lambda(t)X} : Y \mapsto L_{\Lambda(t)X}(Y) := Y\Lambda(t)X + X\Lambda(t)Y. \tag{18}$$

The representation of $L_{\Lambda(t)X}(Y)$ on matrices Y expressed as vectors $Y^{\rm v} \in \mathbb{R}^{s(n)}$ in a basis β of $S(n,\mathbb{R})$ is $D_1F(X,t)\cdot Y^{\rm v}$. Thus the matrix $D_1F(X,t)$ is invertible if and only if $L_{\Lambda(t)X}$ is an invertible map. We will refer to a map of the form

$$L_M : X \mapsto L_M X := XM + MX \tag{19}$$

as a **Lyapunov map** due to its relation to the *Lyapunov equation* $XM + MX = Q$ which arises in the study of the stability of linear control systems. It may be easily verified that a Lyapunov map is linear in X. The map L_M is invertible if no two eigenvalues of M add up to zero (see Callier and Desoer [8], page 138).

Now note that $\Lambda(t)x_*^{\rm m} = x_*^{\rm m}\Lambda(t) = P(t)$ which is positive definite and symmetric. Therefore $L_{\Lambda(t)x_*^{\rm m}}(Y)$ is nonsingular. It follows then that $D_1F(x_*,t)$ is invertible. By the continuity of $D_1F(x,t)$ in x it also follows that $D_1F(x,t)$ remains invertible for all x in a sufficiently small neighborhood of x_*.

4.2 Dynamic Polar Decomposition

The estimator for $D_1F(x_*,t)^{-1}$ will be denoted $\Gamma \in \mathbb{R}^{s(n)\times s(n)}$, so that $D_1F(x_*,t)^{-1} = \Gamma_*$. Using Γ, we may define a dynamic inverse for $F(x,t)$. Let $G : \mathbb{R}^{s(n)} \times \mathbb{R}^{s(n)\times s(n)} \to \mathbb{R}^{s(n)}$; $(w,\Gamma) \mapsto G(w,\Gamma)$ be defined by

$$G(w,\Gamma) := D_1F(x_*,t)^{-1}\big|_{\Gamma_*=\Gamma} \cdot w = \Gamma \cdot w. \tag{20}$$

This makes $G(w,\Gamma)$ a dynamic inverse for $F(x,t) = (x^{\rm m}\Lambda(t)x^{\rm m} - I)^{\rm v}$.

For an estimate of $\dot{x}_*$ we first differentiate $F(x_*,t) = 0$,

$$D_1F(x_*,t)\dot{x}_* + D_2F(x_*,t) = 0 \tag{21}$$

and solve for $\dot{x}_*$,

$$\dot{x}_* = -D_1F(x_*,t)^{-1}D_2F(x_*,t) = -\Gamma_* D_2F(x_*,t). \tag{22}$$

Note that $D_2F(x_*,t) = (x_*^{\rm m}\dot{\Lambda}(t)x_*^{\rm m})^{\rm v}$. Now, substituting x and Γ for x_* and Γ_* we obtain

$$E(x,\Gamma,t) := -\Gamma\left(x^{\rm m}\dot{\Lambda}(t)x^{\rm m}\right)^{\rm v}. \tag{23}$$

To obtain Γ, let $F^\gamma : \mathbb{R}^{s(n)} \times \mathbb{R}^{s(n)\times s(n)} \times \mathbb{R}_+ \to \mathbb{R}^{s(n)\times s(n)}$; $(x,\Gamma,t) \mapsto F^\gamma(x,\Gamma,t)$ be defined by

$$F^\gamma(x,\Gamma,t) := D_1F(x,t)\Gamma - I. \tag{24}$$

A linear dynamic inverse for $F^\gamma(x,\Gamma,t)$ is $G^\gamma : (w,\Gamma) \mapsto G^\gamma[w,\Gamma]$ defined by

$$G^\gamma[w,\Gamma] := \Gamma \cdot w. \tag{25}$$

For an estimator for $\dot{\Gamma}_*$, we differentiate

$$F^\gamma(x_*,\Gamma_*,t) = 0 \tag{26}$$

with respect to t, solve for $\dot{\Gamma}_*$, and substitute x and Γ for x_* and Γ_* respectively to get

$$E^\gamma(x,\Gamma,t) := -\Gamma\left(\frac{d}{dt}D_1F(x,t)\right)\bigg|_{\dot{x}_*=E(x,\Gamma,t)} \cdot \Gamma. \tag{27}$$

Combining the E's, F's, and G's from (23), (16), (20), (27), (24), and (25), we obtain the dynamic inverter

$$\left\{\begin{array}{l} \dot{x} = -\mu\Gamma\left(x^{\rm m}\Lambda(t)x^{\rm m} - I\right)^{\rm v} - \Gamma\left(x^{\rm m}\dot{\Lambda}(t)x^{\rm m}\right)^{\rm v} \\ \dot{\Gamma} = -\mu\Gamma\left(D_1F(x,t)\Gamma - I\right) \\ \qquad -\Gamma\left(\frac{d}{dt}D_1F(x,t)\right)\big|_{\dot{x}_*=E(x,\Gamma,t)}\cdot\Gamma \end{array}\right. \tag{28}$$

Initial conditions for the dynamic inverter may be set so that

$$x(0) \approx (P(0)^{-1})^{\rm v} \quad \text{and} \quad \Gamma(0) \approx D_1F((P(0)^{-1})^{\rm v},t)^{-1}. \tag{29}$$

Combining the results above with the dynamic inversion theorem, Theorem 3.1 of [1] gives the following theorem.

Theorem 4.1 *Let $A(t)$ be in $GL(n,\mathbb{R})$ for all $t \in \mathbb{R}_+$. Let the polar decomposition of $A(t)$ be $A(t) = P(t)U(t)$ with $P(t) \in S(n,\mathbb{R})$ the positive definite symmetric square root of $\Lambda(t) := A(t)A(t)^T$ and $U(t) \in O(n,\mathbb{R})$ for all $t \in \mathbb{R}_+$. Let x be in $\mathbb{R}^{s(n)}$, and let Γ be in $\mathbb{R}^{s(n)\times s(n)}$. Let $(x(t),\Gamma(t))$ denote the solution of the dynamic inverter (28) where $F(x,t)$ is given by (16). Then there exists a $\tilde{\mu}$ such that if $\mu > \tilde{\mu}$, and $(x(0),\Gamma(0))$ is sufficiently close to $((P(0)^{-1})^{\rm v}, D_1F((P(0)^{-1})^{\rm v}$ then*

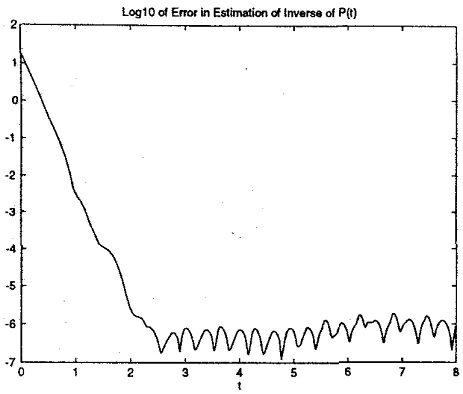

Figure 1: The error $\log_{10}(\|x(t)^{\mathrm{m}}\Lambda(t)x(t)^{\mathrm{m}} - I\|_{\infty})$.

1. $\Lambda(t)(x(t))^{\mathrm{m}}$ *exponentially converges to* $P(t)$,
2. $(x(t))^{\mathrm{m}}A(t)$ *exponentially converges to* $U(t)$, *and*
3. $A(t)((x(t))^{\mathrm{m}})^2$ *exponentially converges to* $A(t)^{-1}$.

$\diamond$

An example of the polar decomposition of a 2×2 matrix will illustrate our results.

Example 4.2 Let

$$A(t) := \begin{bmatrix} 10 + \sin(10t) & \cos(t) \\ -t & 1 \end{bmatrix}. \qquad (30)$$

Dynamic inversion of $A(t)$ using (28) was simulated over the interval $t \in [0, 8]$ using the adaptive step-size Runge-Kutta integrator **ode45** from Matlab, with the default tolerance of 10^{-6}. Initial conditions were set to be

$$x(0) = \Lambda^{1/2}(0)^{\mathrm{v}} + e_x, \qquad \Gamma(0) = D_1F(x(0), t)^{-1} \qquad (31)$$

where $e_x = [-0.55, 0.04, -2.48]^T$ is an error that has been deliberately added to demonstrate the error transient of the dynamic inverter. The value of μ was set to 10. For more details of this example see [2].

Figure 1 shows $\log_{10}(\|x(t)^{\mathrm{m}}\Lambda(t)x(t)^{\mathrm{m}} - I\|_{\infty})$ indicating the extent to which x^{m}, the estimator for $P(t)^{-1}$ fails to be the square root of $\Lambda(t) = A(t)A(t)^T$.

$\triangle$

Remark 4.3 It is interesting to note that $P(t)^{-1}$, besides being a solution to $x^{\mathrm{m}}\Lambda(t)x^{\mathrm{m}} - I = 0$ is also a solution to $\Lambda(t)(x^{\mathrm{m}})^2 - I = 0$ as well as $(x^{\mathrm{m}})^2\Lambda(t) - I = 0$. But $\Lambda(t)(x^{\mathrm{m}})^2 - I$ and $(x^{\mathrm{m}})^2\Lambda(t) - I$ are not, in general, symmetric even when $\Lambda(t)$ and x^{m} are symmetric. Though exponential convergence is still guaranteed when using these forms, the flow of Γ is not confined to $S(n, \mathbb{R})$. Using these forms would increase the number of equations in the dynamic inverter by $n(n-1)/2 + n^2 - s(n)^2$ since, not only would the right hand side of the top equation of (28) no longer be symmetric, but Γ would be $n^2 \times n^2$ rather than $s(n) \times s(n)$. $\triangle$

5. Polar Decomposition and Inversion of Fixed Matrices

In the dynamic inversion techniques of Sections 2 and 4 we assumed that we had available an approximation of $A^{-1}(0)$ with which to set $\Gamma(0)$ in the dynamic inversion of $A(t)$. Thus we would need to invert at least one fixed matrix, $A(0)$, in order to start the dynamic inverter. Methods of fixed matrix inversion presented in Section 3 had the potential disadvantage of either producing exact inversion only asymptotically as $t \to \infty$, or of only working on matrices with no eigenvalues in $(-\infty, 0) = \phi$. The question naturally arises then, how might we use dynamic inversion to invert *any* fixed matrix so that the exact inverse is available by a fixed time. In this section, by appealing to both homotopy and polar decomposition, we give an answer to this question.

Let M be in $GL(n, \mathbb{R})$ with $P = P^T > 0$, $UU^T - I$, and $M = PU$. Helmke and Moore (see [6], pages 150-152) have described a gradient flow (using the Riemannian metric on $\mathbb{R}^{n \times n}$) for the function $\|A - UP\|$,

$$\begin{aligned} \dot{U} &= UPM^TU - MP \\ \dot{P} &= -2P + M^TU + U^TM \end{aligned} \qquad (32)$$

where P is meant to approximate P and U is meant to approximate U. Asymptotically, this system produces products P_* and U_* satisfying $A - P_*U_* = 0$ for almost all initial conditions as $t \to \infty$. A difficulty with this approach, as the authors point out, is that positive definiteness of the approximator P is not guaranteed. The method we describe in this section provides polar decomposition of any nonsingular matrix in finite time, with the positiveness of P guaranteed. It will be seen that our method does not rely upon a gradient structure.

Let $\Lambda(t) := (1-t)I + tMM^T$ so that $\Lambda(0) = I$, $\Lambda(1) = MM^T$, and for all $t \in [0, 1]$, $\Lambda(t)$ is positive definite and symmetric. Let $P_\Lambda(t)$ denote the positive definite symmetric square root of $\Lambda(t)$. Let the estimator of $P_\Lambda^{-1}(t)$ be $x^{\mathrm{m}} \in \mathbb{R}^{n \times n}$. For this definition of $\Lambda(t)$ we have $\dot{\Lambda}(t) = MM^T - I$. Now we may apply the dynamic inverter of Section 4 in order to perform the polar decomposition of M. By inspection we see that $x_*(0) = I^{\mathrm{v}}$ and $\Gamma_*(0) = \frac{1}{2}I^{\mathrm{v}}$. If we set $\Gamma(0) = I^{\mathrm{v}}$ and $\Gamma(0) = \frac{1}{2}I^{\mathrm{v}}$, then the dynamic inversion theorem, Theorem 3.1 of [1] and the results of the last section assure us that $x(t)^{\mathrm{m}} \equiv P_\Lambda(t)^{-1}$ for all $t \geq 0$, and thus $x(1)^{\rm m} = P_\Lambda(1)^{-1}$. Consequently

$$\begin{array}{ll} x(1)^{\rm m} = P^{-1}, & MM^Tx(1)^{\rm m} = P, \\ x(1)^{\rm m}M = U, & M^T(x(1)^{\rm m})^2 = M^{-1}. \end{array} \tag{33}$$

Note that $\dot{\Lambda}(t) = MM^T = 0$ if and only if M is unitary, in which case $M^{-1} = M^T$.

Combining the results of this section with the results of the last section gives the following Theorem.

Theorem 5.1 *Let M be in $GL(n,\mathbb{R})$. Let the polar decomposition of M be $M = PU$ with $P \in S(n,\mathbb{R})$ the positive definite symmetric square root of MM^T and $U \in O(n,\mathbb{R})$. Let x be in $\mathbb{R}^{s(n)}$, and let Γ be in $\mathbb{R}^{s(n)\times s(n)}$. Let $x(0) = I^{\rm v}$ and $\Gamma(0) = \frac{1}{2}I$. Let $(x(t), \Gamma(t))$ denote the solution of*

$$\left\{ \begin{array}{l} \dot{x} = -\mu G(F(x,t),\Gamma) + E(x,\Gamma) \\ \dot{\Gamma} = -\mu G^\gamma[F^\gamma(\Gamma,x)] + E^\gamma(x,\Gamma). \end{array} \right. \tag{34}$$

where

$$\begin{array}{l} \Lambda(t) = (1-t)I + tMM^T \\ F(x,t) = x^{\rm m}\Lambda(t)x^{\rm m} - I \\ G(w,\Gamma) = \Gamma \cdot w \\ E(x,\Gamma) = -\Gamma(x^{\rm m}(MM^T - I)x^{\rm m})^{\rm v} \\ G^\gamma[w,\Gamma] = \Gamma \cdot w \\ F^\gamma(x,\Gamma,t) = D_1F(x,t)\Gamma - I \\ E^\gamma(x,\Gamma) = -\Gamma \left(\frac{d}{dt}D_1F(x,t)\right)\big|_{\dot{x}=E(x,\Gamma)} \cdot \Gamma \end{array} \tag{35}$$

Then for any $\mu > 0$, $MM^Tx(1) = P$, $x(1)M = U$, and $Mx(1)^2 = M^{-1}$. ◊

Remark 5.2 As in Theorem 3.4 we can force $\Gamma^{\rm m}$ to equal P^{-1} at any time $t_1 > 0$ by substituting t/t_1 for t in $\Lambda(t)$, and proceeding with the derivation of the dynamic inverter as above. Then $\dot{\Lambda} = \frac{1}{t_1}(MM^T - I)$ and $x^{\rm m}(t_1) = P^{-1}$.

Example 5.3 A digital computer simulation of a dynamic inverter for the polar decomposition of a fixed 2×2 matrix was performed. The integration was done in Matlab [9] using `ode45` an adaptive step size Runge-Kutta routine using the default tolerance of 10^{-6}. The matrix M was chosen randomly to be

$$M = \left[\begin{array}{cc} 7 & -3 \\ -24 & -3 \end{array} \right]. \tag{36}$$

The value of μ was set to 10.

The final value ($t = 1$) of the error $\|x(t)^{\rm m}MM^Tx(t)^{\rm m} - I\|_\infty$ was $\|x(1)^{\rm m}\Lambda(1)x(1)^{\rm m} - I\|_\infty = 6.9575 \times 10^{-7}$. Final values of P, U, and A^{-1} were

$$\begin{array}{l} P = MM^Tx(1)^{\rm m} = \left[\begin{array}{cc} 5.2444 & -5.5223 \\ -5.5223 & 23.5479 \end{array} \right] \\ U = x(1)^{\rm m}M = \left[\begin{array}{cc} 0.3473 & -0.9377 \\ -0.9377 & -0.3473 \end{array} \right] \\ M^{-1} = M^T(x(1)^{\rm m})^2 = \left[\begin{array}{cc} 0.0323 & -0.0323 \\ -0.2581 & -0.0753 \end{array} \right] \end{array}$$

△

6. Summary

We have seen how the polar decomposition and inversion of time varying and fixed matrices may be accomplished by continuous-time dynamical systems. Our results are easily modified to provide solutions for time varying and fixed linear equations of the form $A(t)x = b$.

Standard discrete matrix inversion routines do not take advantage of one's knowledge of $\dot{A}(t)$. Dynamic inversion, on the other hand, by utilizing derivative estimation based upon such knowledge, may lead to increases in computational efficiency. We have also seen that dynamic inversion in the matrix context provides a useful and general conceptual framework through which to view other methods of dynamic computation such as neural networks.

Acknowledgements. The authors are grateful to C.A. Desoer for his comments and advice.

7. References

[1] N. H. Getz and J. E. Marsden, "A dynamic inverse for nonlinear maps," in *34th IEEE Conference on Decision and Control*, (New Orleans), IEEE, 13-15 December 1995.

[2] N. H. Getz and J. E. Marsden, "Dynamical methods for polar decomposition and inversion of matrices," Tech. Rep. 624, Center for Pure and Applied Mathematics, Berkeley, California, 5 January 1995. Submitted to Linear Algebra and its Applications.

[3] N. H. Getz, *Dynamic Inversion of Nonlinear Maps with Applications to Nonlinear Control.* PhD thesis, University of California at Berkeley, October 1995.

[4] J.-S. Jang, S.-Y. Lee, and S.-Y. Shin, "An optimization network for matrix inversion," in *Neural Information Processing Systems* (D. Z. Anderson, ed.), pp. 397-401, New York: American Institute of Physics, 1988.

[5] J. Wang, "A recurrent neural network for real-time matrix inversion," *Applied Mathematics and Computation*, vol. 55, pp. 89–100, 1993.

[6] U. Helmke and J. B. Moore, *Optimization and Dynamical Systems.* Communications and Control Engineering, New York: Springer-Verlag, 1994.

[7] R. A. Horn and C. R. Johnson, *Matrix Analysis.* New York: Cambridge University Press, 1985.

[8] F. Callier and C. Desoer, *Linear System Theory.* Springer Texts in Electrical Engineering, New York: Springer-Verlag, 1991.

[9] *Matlab.* Natick, Mass.: The MathWorks, Inc., 1992.