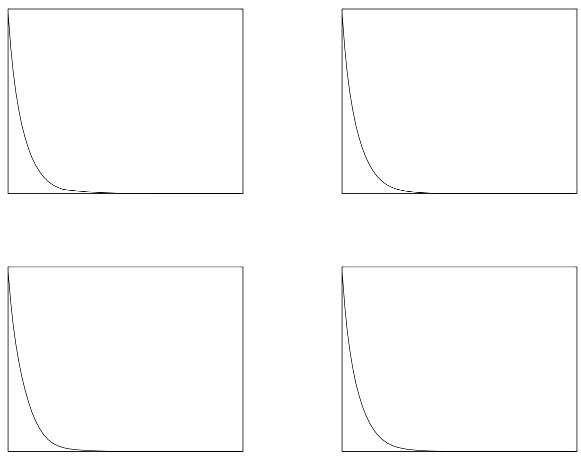
DYNAMICAL METHODS FOR POLAR DECOMPOSITION AND INVERSION OF MATRICES



Neil H. Getz and Jerrold E. Marsden

Memorandum No. CPAM 629

5 January 1995

CENTER FOR PURE AND APPLIED MATHEMATICS

Department of Mathematics
University of California at Berkeley
Berkeley, CA 94720

Dynamical Methods for Polar Decomposition and Inversion of Matrices*

Neil H. Getz
Department of Electrical Engineering and Computer Sciences
University of California at Berkeley
Berkeley, CA 94720
getz@eecs.berkeley.edu

Jerrold E. Marsden
Control and Dynamical Systems 104-44
California Institute of Technology
Pasadena, CA 91125
marsden@cds.caltech.edu

January 5, 1995

Abstract

We show how one may obtain the polar decomposition as well as regular and generalized inversion of fixed and time-varying matrices using a class of dynamical systems. First a dynamical system is constructed that causes an initial approximation of a matrix inverse to flow exponentially toward the true time-varying inverse. The same dynamical method may be applied to the inversion of fixed matrices. By appealing to a time-parameterized homotopy from the identity to a fixed matrix, and applying the first result on the inversion of time-varying matrices, we show how any positive-definite fixed matrix may be dynamically inverted without an initial guess at the inverse, the exact inverse being produced in finite time. We then construct a dynamical system that solves for the polar decomposition components of a time varying matrix given an initial approximation for the inverse of the positive-definite symmetric part of the polar decomposition. As a byproduct, this method gives another method of inverting time-varying matrices. Finally by using homotopy again, we show how the dynamic polar decomposition may be applied to fixed matrices allowing us to dynamically invert any fixed matrix in finite time. Our approach is metric free and offers an alternative and some improvement over extant polar decomposition methods based on gradient flows. All results are derived within the framework of our previous work on the dynamic inversion of nonlinear maps.

*This is Center for Pure and Applied Mathematics Memorandum 629. Revised November 27, 1995.

Keywords: dynamical systems, polar decomposition, dynamic inversion, matrix inversion, neural networks, homotopy, gradient flow, optimization, inverse problems.

1 Introduction

We have recently introduced in [1, 2] a methodology for the construction of dynamic systems which generate an approximations $\theta(t)$ to solutions $\theta_*(t)$ of nonlinear vector equations of the form $F(\theta,t) = 0$. We call this technique *dynamic inversion* since we view it as a dynamic method for the solution of finite-dimensional time-varying inverse problems. In dynamic inversion, one supplies an initial condition $\theta(0)$ sufficiently close to $\theta_*(0)$. A dynamic inverter then generates $\theta(t)$ which converges to $\theta_*(t)$ exponentially.

Dynamic inversion has evolved out of the need to solve and track solutions to time-varying finite-dimensional inverse problems in the context of the automatic control of nonlinear systems. As a side benefit its region of attraction has been shown [1, 2] to be a superset of the region of attraction of Newton's method and its variants since it is not dissuaded from convergence by local minima and maxima, and does not require smoothness of $F(\theta,t)$. Dynamic inversion also allows one to incorporate knowledge of how a root varies in time into the root-finding procedure.

One may also pose inverses and square roots of time-varying matrices as a solutions to equations of the form $F(\theta,t) = 0$. Motivated by this realization, in this paper we will use dynamic inversion to construct dynamic systems that perform polar decomposition as well as matrix inversion.

1.1 Notation

Here, for easy reference, we define some symbols and notation that we will use in this paper.

For any integer $k > 0$, let $\underline{k} := \{1, 2, \ldots, k\}$.

The l_2 norm on $A = [A_{i,j}]_{i,j\in\underline{n}} \in \mathbb{R}^{n\times n}$, is defined by

$$||A||_2 := (\sum_{i,j\in\underline{n}} |A_{i,j}|^2)^{1/2}. \tag{1.1}$$

The l_∞ norm on A is defined by $||A||_\infty := \max_{i,j\in\underline{n}} |A_{i,j}|$.

The symbol $GL(n)$ denotes the group of nonsingular n-by-n matrices. The symbol $O(n)$ denotes the group of orthogonal n-by-n matrices. The symbol $S(n)$ denotes the group of symmetric n-by-n matrices. For convenience we define $s(n)$ to be the dimension of $S(n)$, i.e. $s(n) := n(n+1)/2$.

Given a Banach space $\mathcal{S}$, and a positive real number r, define the ball $\mathcal{B}_r := \{z \in \mathcal{S} \mid ||z||_2 \leq r\}$. The particular set $\mathcal{S}$ will be clear from context.

Let $D_i F(a_1, \ldots, a_n)$ denote the derivative of a map $F(a_1, \ldots, a_n)$ with respect to a_i. The k^{th} derivative of $F(a_1, \ldots, a_n)$ with respect to a_i is denoted $D_i^k F(a_1, \ldots, a_n)$.

The positive real numbers are denoted $\mathbb{R}_+ := \{\lambda \in \mathbb{R} | \lambda \geq 0\}$.

The *spectrum* of $M \in GL(n)$ is the set of eigenvalues of M and is denoted $\sigma(M)$.

1.2 Previous Work

Continuous-time dynamic methods of solving matrix equations have appeared before. Indeed, any dynamic system on a matix space for which an asymptotically stable equilibrium exists may be considered to be a dynamic inverter that solves for its equilibrium. Continuous-time dynamic methods for determining eigenvalues date back at least as far as Rutishauser [3, 4].

Recently, however, there has been a great renewal of interest in the use of continuous-time dynamical systems for computation. Though the most prominent component of this renewed interest has been the activity and interest surrounding artificial neural networks, profound insights have been generated in other directions as well. Brockett [5] has recently shown how one can use matrix differential equations to perform computation often thought of as being intrinsically discrete: the sorting of lists, matrix diagonalization, and the solving of linear programming problems. Chu [6] has studied the Toda flow as a continuous-time analog of the QR algorithm. Bloch [7, 8] has shown how Hamiltonian systems may be used to solve principal component and linear programming problems. Chu [9] and Chu and Driessel [10] have explored the use of differential equations in solving linear algebra problems. Smith [11], and Helmke et al. [12] have constructed dynamical systems that perform singular-value decomposition. Dynamic methods of matrix inversion have also appeared in the artificial neural network literature. See for instance Jang et al. [13] and Wang [14]. For a review of dynamic matrix methods as well as a comprehensive list of references for dynamic approaches to optimization see [15].

A dynamic decomposition related to polar decomposition of fixed matrices has also appeared in Helmke and Moore [15], though, as the authors point out, their gradient based method does not guarantee the positive definiteness of the symmetric component of the polar decomposition. Using dynamic inversion we will derive a system that produces the desired inverse and polar decomposition products at any fixed time $t_1 > 0$ with guaranteed positive-definiteness of the symmetric component.

As far as we know, all prior continuous-time dynamic approaches to inversion of matrix equations use gradient flows. In contrast, dynamic inversion, as we formulate it, does not require the requisite metric needed to define a gradient. Though we will see in Section 4.1 that gradient approaches fit well into the dynamic inversion framework, the main results of the present paper do not rely upon a metric structure.

1.3 Main Results

The main results of this chapter are as follows: We will construct dynamic systems that

1. invert time-dependent matrices asymptotically,

2. invert a spectrally restricted matrix by a prescribed time,

3. invert and decompose any t-dependent invertible matrix into its polar decomposition factors,

4. invert and decompose any constant nonsingular matrix into its polar decomposition factors by a prescribed time.

1.4 Overview

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 3 we summarize some results of [1] on the inversion of time-varying matrices and show some further applications of time-varying matrix inversion. Motivated by the desire to obtain the initial inverse dynamically, in Section 4 we will consider the problem of inverting constant matrices. By using a matrix homotopy from the identity we will use the results of Section 3 to produce exact inversion of a restricted class of constant matrices, including positive-definite matrices, by a prescribed time. In Section 5 we will consider the polar decomposition of a time-varying matrix. We will show how, starting from a good guess at the initial value of the inverse of the positive-definite part of the polar decomposition, we may construct a dynamic system that produces an exponentially convergent estimate of the inverse of the positive-definite symmetric part. From this estimate and the original matrix we may obtain the decomposition products as well as the inverse. Then in Section 6 we revisit the problem of constant matrix inversion and show how, combining homotopy with dynamic polar decomposition, we may dynamically produce the polar decomposition factors as well as the inverse of *any* constant matrix by a prescribed time *without* requiring an initial guess.

2 A Review of Dynamic Inversion

Dynamic inversion [1, 2], is a method of using dynamics to provide an estimate of time-varying roots of time-dependent maps. Alternatively it may be regarded as a framework in which to view a number of methods for the inversion of nonlinear maps. In dynamic inversion one associates with a map $F(\theta, t)$ a dynamical system $\dot{\theta} = \Phi(\theta, t)$ with the crucial property that a continuous isolated solution $\theta_*(t)$ of $F(\theta, t) = 0$ is exponentially attractive[1]. Dynamic inversion depends intimately upon the notion of a *dynamic inverse* whose definition we now recall.

[1]See [1, 2] for another version of dynamic inversion in which an arbitrarily small neighborhood of $\theta_*(t)$ is exponentially attractive. This is particularly useful when $F(\theta, t)$ is not differentiable.

Definition 2.1 For $F : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}^n; (\theta, t) \mapsto F(\theta, t)$ let $\theta_*(t)$ be a continuous isolated solution of $F(\theta, t) = 0$. A map $G : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}^n; (w, t) \mapsto G[w, \theta, t]$ is called a **dynamic inverse** of $F(\theta, t)$ on the ball $\mathcal{B}_r := \{z \in \mathbb{R}^n | \, \|z\| \leq r\}$, $r > 0$, if

1. the map $G[F(\theta, t), \theta, t]$ is Lipschitz in θ, piecewise-continuous in t, and
2. there is a real constant β, with $0 < \beta < \infty$, such that

Dynamic Inverse Criterion

$$z^T G\left[F\left(z + \theta_*(t), t\right), z + \theta_*(t), t\right] \geq \beta \|z\|_2^2 \tag{2.1}$$

for all $z \in \mathcal{B}_r$.

▲

See [1, 2] for properties of dynamic inverses.

In this paper we will use estimates of $D_1 F(\theta_*(t), t)^{-1}$ to form linear dynamic inverses, e.g. $G[w, \theta] = D_1 F(\theta, t) \cdot w$ where θ is close to $\theta_*(t)$. In some cases an estimate of $D_1 F(\theta_*(t), t)^{-1}$ will be determined dynamically. We will also have occasion to use the dynamic inverse $G[w, \theta] = D_1 F(\theta_*, t)^T \cdot w$.

Sufficient conditions for the existence of a linear dynamic inverse for all $t \in \mathbb{R}_+$ are mild as indicated by the following lemma.

Lemma 2.2 Sufficient Conditions for Existence of a Dynamic Inverse. *For $F : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}^n$; $(\theta, t) \mapsto F(\theta, t)$, let $\theta_*(t)$ be a continuous isolated solution of $F(\theta, t) = 0$. Let $F(\theta, t)$ be C^2 in θ and continuous in t. Assume that the following are true:*

1. *$D_1 F(\theta_*(t), t)$ is nonsingular for all t;*
2. *$D_1 F(\theta_*(t), t)$ and $D_1 F(\theta_*(t), t)^{-1}$ are bounded uniformly in t;*
3. *for all $z \in \mathcal{B}_r$, $D_1^2 F(z + \theta_*(t), t)$ is bounded uniformly in t.*

Under these conditions there exists an $r > 0$ independent of t, and a function $G : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}^n$, $(w, \theta, t) \mapsto G[w, \theta, t]$ such that for each $t > 0$ and for all θ satisfying $\theta - \theta_*(t) \in \mathcal{B}_r$, $G[w, \theta, t]$ is a dynamic inverse of $F(\theta, t)$. ◆

Proof of Lemma 2.2: See [1] [2]. ■

The dynamic inversion theorem, Theorem 2.3 below, ties the dynamic inverse of a map $F(\theta, t)$ to dynamical estimation of the solution θ_* of $F(\theta, t) = 0$.

Theorem 2.3 Dynamic Inversion Theorem – Vanishing Error. *Let $\theta_*(t)$ be a continuous isolated solution of $F(\theta, t) = 0$, with $F : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}^n$;*

$(\theta,t)\mapsto F(\theta,t)$. Assume that $G:\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}_+\to\mathbb{R}^n$; $(w,\theta,t)\mapsto G[w,\theta,t]$, is a dynamic inverse of $F(\theta,t)$ on $\mathcal{B}_r$, for some finite $\beta>0$. Let $E:\mathbb{R}^n\times\mathbb{R}_+\to\mathbb{R}^n$; $(\theta,t)\mapsto E(\theta,t)$ be locally Lipschitz in θ and continuous in t. Assume that for some constant $\kappa\in(0,\infty)$, $E(\theta,t)$ satisfies

$$\left\| E\left(z+\theta_*(t),t\right)-\dot{\theta}_*(t)\right\|_2 \leq \kappa\|z\|_2 \tag{2.2}$$

for all $z\in\mathcal{B}_r$. Let $\theta(t)$ denote the solution to the system

$$\dot{\theta} = -\mu\, G\left[F\left(\theta,t\right),\theta,t\right]+E(\theta,t) \tag{2.3}$$

with initial condition $\theta(0)$ satisfying $\theta(0)-\theta_*(0)\in\mathcal{B}_r$. Then

$$\|\theta(t)-\theta_*(t)\|_2 \leq \|\theta(0)-\theta_*(0)\|_2\, e^{-(\mu\beta-\kappa)t} \tag{2.4}$$

for all $t\in\mathbb{R}_+$, and in particular if $\mu>\kappa/\beta$, then $\theta(t)$ converges to $\theta_*(t)$ exponentially as $t\to\infty$. ♦

Proof of Theorem 2.3: See [1, 2]. ■

Remark 2.4 Dynamic Inversion with Perfect Initial Conditions. If $\theta(0)=\theta_*(0)$, then the conditions of Theorem 2.3 guarantee that $\theta(t)\equiv\theta_*(t)$ for all $t\in\mathbb{R}_+$. So in a sense, we need only solve the inverse problem at a single instant $t=0$. Then the dynamic inversion takes care of maintaining the inversion for all t. In many applications, however, one can only expect $\theta(0)\approx\theta_*(0)$. ▲

Theorem 2.5 Dynamic Inversion with Dynamic Determination of a Dynamic Inverse. *Let $F(\theta,t)$ satisfy the assumptions of Lemma 2.2. Then for $\Gamma(0)$ sufficiently close to $D_1F(\theta_*,0)^{-1}$, and $\theta(0)$ sufficiently close to $\theta_*(0)$, the solution $(\Gamma(t),\theta(t))$ of*

$$\boxed{\begin{aligned}\begin{bmatrix}\dot{\Gamma}\\ \dot{\theta}\end{bmatrix} &= -\mu\begin{bmatrix}\Gamma & 0\\ 0 & \Gamma\end{bmatrix}\begin{bmatrix}D_1F(\theta,t)\Gamma-I\\ F(\theta,t)\end{bmatrix}\\ &\quad+\begin{bmatrix}-\Gamma\,\frac{d}{dt}D_1F(\theta,t)\big|_{\dot{\theta}=-\Gamma D_2F(\theta,t)}\,\Gamma\\ -\Gamma D_2F(\theta,t)\end{bmatrix}\end{aligned}} \tag{2.5}$$

satisfies $(\Gamma(t),\theta(t))\to(D_1F(\theta_*,t)^{-1},\theta_*(t))$ as $t\to\infty$. Furthermore, for sufficiently large $\mu>0$, the convergence is exponential, i.e. there exist $k_1>0$ and $k_2>0$ such that

$$\|(\Gamma(t),\theta(t))-(\Gamma_*(t),\theta_*(t))\|_2 \leq k_1\|(\Gamma(0),\theta(0))-(\Gamma_*(0),\theta_*(0))\|_2 e^{-k_2t} \tag{2.6}$$

for all $t\geq0$, where $\Gamma_*(t)=D_1F(\theta_*(t),t)^{-1}$. ♦

In the present paper we will assume that a closed form for $F(\theta, t)$ is available and is C^1 in θ, though in general $F(\theta, t)$ need not have a closed form or be differentiable to have a dynamic inverse. Under suitable conditions of differentiability a θ- and t-dependent estimator for $\dot{\theta}_*$, $E(\theta, t)$, may be obtained by differentiating $F(\theta_*(t), t) = 0$ with respect to t, and substituting θ for θ_*. Such an estimator becomes arbitrarily precise as θ approaches θ_*. Approximations of $E(\theta, t)$ may also possess this property.

3 Inverting Time-Varying Matrices

Consider the problem of estimating the inverse $\Gamma_*(t) \in \mathbb{R}^{n\times n}$ of a time-varying matrix $A(t) \in GL(n, \mathbb{R})$, where $GL(n, \mathbb{R})$ denotes the group of invertible matrices in $\mathbb{R}^{n\times n}$. Assume that we have representations for both $A(t)$ and $\dot{A}(t)$, and that $A(t)$ is C^1 in t. Let Γ be an element of $\mathbb{R}^{n\times n}$.

In order for Γ_* to be the inverse of $A(t)$, Γ_* must satisfy

$$A(t)\Gamma - I = 0 \tag{3.1}$$

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

$$F(\Gamma, t) := A(t)\Gamma - I \tag{3.2}$$

We will refer to the solution of $F(\Gamma, t) = 0$ as $\Gamma_*(t)$. To obtain an estimator $E(\Gamma, t)$ for $\dot{\Gamma}_*(t)$, differentiate $A\Gamma_* = I$ with respect to t, solve the resulting expression for $\dot{\Gamma}_*$, replace A^{-1} by θ_*, and then replace Γ_* by Γ in the resulting expression to get

$$E(\Gamma, t) := -\Gamma\dot{A}(t)\Gamma \tag{3.3}$$

Differentiate $F(\Gamma, t)$ with respect to Γ to get

$$D_1F(\Gamma, t) = A(t) \tag{3.4}$$

whose inverse is Γ_*. So a choice of dynamic inverse is

$$G[w, \Gamma, t] := \Gamma \cdot w \tag{3.5}$$

for Γ sufficiently close to $\Gamma_* = A^{-1}(t)$ and with $w \in \mathbb{R}^{n\times n}$. The dynamic inverter for this problem then takes the form

$$\dot{\Gamma} = -\mu G[F(\Gamma, t), \Gamma] + E(\Gamma, t) \tag{3.6}$$

or, expanded,

$$\dot{\Gamma} = -\mu\Gamma(A(t)\Gamma - I) - \Gamma\dot{A}(t)\Gamma \tag{3.7}$$

and we choose as initial conditions $\Gamma(0) \approx \Gamma_*(0) = A^{-1}(0)$ so that the estimation error starts small. Theorem 2.3 guarantees that for sufficiently large μ, and for $\Gamma(0)$ sufficiently close to $A^{-1}(0)$, equation (3.7) will produce an estimator $\Gamma(t)$ whose error $\|\Gamma(t) - \Gamma_*(t)\|$ decays exponentially to zero at a rate determined by our choice of μ.

There are an infinite number of dynamic inverses $G[w, \Gamma, t]$ which would suffice for the dynamic inversion of $F(\Gamma, t)$ (3.2). One convenient alternative is to use $G[w, t] = A(t)^T \cdot w$. The resulting dynamic inverter appears in Example 3.2 below.

We summarize the results above into the following theorem:

Theorem 3.1 Dynamic Inversion of Time-Varying Matrices. *Let $A(t) \in GL(n, \mathbb{R})$ be C^1 in t, with $A(t)$, $A(t)^{-1}$, and $\dot{A}(t)$ bounded on $[0, \infty)$. Let $G[w, \Gamma, t]$ be a dynamic inverse of $F(\Gamma, t) = A(t)\Gamma - I$ for all $t \in \mathbb{R}_+$, and for all Γ such that $\Gamma - \Gamma_*$ is in $\mathcal{B}_r$. Let $\Gamma(t) \in \mathbb{R}^{n\times n}$ be the solution to*

$$\dot{\Gamma} = -\mu G[A(t)\Gamma - I, \Gamma, t] - \Gamma \dot{A}(t) \Gamma \tag{3.8}$$

with $\|\Gamma(0) - \Gamma_*(0)\| \leq r < \infty$. Then for sufficiently small r, there exists a $\tilde{\mu} > 0$ and a $k > 0$ such that for all $\mu > \tilde{\mu}$, and for all $t \geq 0$,

$$\|\Gamma(t) - \Gamma_*(t)\|_2 \leq r e^{-kt} \tag{3.9}$$

In particular $\lim_{t\to\infty} \Gamma(t) = A(t)^{-1}$. ♦

Example 3.2 A Dynamic Inverter for a Time-Varying Matrix. Let

$$G[w, t] := A(t)^T \cdot w \tag{3.10}$$

Then for sufficiently large constant $\mu > 0$, and for $\Gamma(0)$ sufficiently close to $A(0)^{-1}$, the solution $\Gamma(t)$ of

Dynamic Inverter for a Time-Varying Matrix

$$\dot{\Gamma} = -\mu A(t)^T \left(A(t)\Gamma - I\right) - \Gamma \dot{A}(t) \Gamma \tag{3.11}$$

approaches $A(t)^{-1}$ exponentially as $t \to \infty$. ▲

The dynamic inverse $G[w, \Gamma] = \Gamma \cdot w$ may also be used instead of $G[w, t] = A(t)^T \cdot w$ to get

$$\dot{\Gamma} = -\mu \Gamma \left(A(t)\Gamma - I\right) - \Gamma \dot{A}(t) \Gamma \tag{3.12}$$

as another dynamic inverter for time-varying matrices.

Example 3.3 Dynamic Inversion of a Mass Matrix. Consider a finite dimensional mechanical system modeled by the implicit second order differential equation

$$M(q)\ddot{q} + N(q,\dot{q}) = 0 \tag{3.13}$$

Usually the matrix $M(q)$ is positive-definite and symmetric for all q since the kinetic energy, $(1/2)\dot{q}^T M(q)\dot{q}$, is greater than zero for all $\dot{q} > 0$. It is often convenient to express such systems in an explicit form, with $\ddot{q}$ alone on the left side of a second order ordinary differential equation. To do so we will invert $M(q)$ dynamically.

Let Γ be a symmetric estimator for $M(q)^{-1}$. 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 3.1. Then the system

Dynamic Inverter for a Mass Matrix

$$\left\{ \begin{array}{rcl} \dot{\Gamma} & = & -\mu\Gamma\left(M(q)\Gamma - I\right) - \Gamma\left[\frac{\partial M_{i,j}(q)}{\partial q}\dot{q}\right]_{i,j\in\underline{n}} \cdot \Gamma \\ \ddot{q} & = & \Gamma N(q,\dot{q}) \end{array} \right. \tag{3.14}$$

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

Remark 3.4 Symmetry and the Choice of Dynamic Inverse. In Example 3.3, $M(q)$ is symmetric, as is its inverse $M(q)^{-1}$. The right hand side of (3.14) is also symmetric, hence if $\Gamma(0)$ is symmetric, so will be $\Gamma(t)$ for all t. If we had chosen $G[w,q] := M(q)^T \cdot w$ as a dynamic inverse (see, for instance, Example 3.2) we would not have had this symmetry. The symmetry allows us to cast the top equation of (3.14) on the space $S(n,\mathbb{R})$ of symmetric $n \times n$ matrices thereby reducing the complexity of the dynamic inverter with respect to the nonsymmetric case; what would otherwise be n^2 equations (3.14) is reduced to $s(n) := n(n+1)/2$ equations. ▲

3.1 Left and Right Inversion of Time-Varying Matrices

Consider a matrix $A(t) \in \mathbb{R}^{m\times n}$. Assume that $A(t)$ is of full rank for all $t \geq 0$. We consider two cases: (1) If $m \leq n$, then $A(t)$ has a right inverse $\Gamma_*(t) \in \mathbb{R}^{n\times m}$ satisfying

$$F(\Gamma,t) := A(t)\Gamma - I = 0 \tag{3.15}$$

It is easily verified that

$$G[w] := \Gamma \cdot w \tag{3.16}$$

is a dynamic inverse for $F(\Gamma, t)$ for Γ sufficiently close to $\Gamma_* = A(t)^T(A(t)A(t)^T)^{-1}$. Differentiate $F(\Gamma_*, t) = 0$ with respect to t, solve for $\dot{\Gamma}_*$, and replace Γ_* by Γ to get the derivative estimator

$$E(\Gamma, t) := -\Gamma \dot{A}(t)\Gamma \tag{3.17}$$

Thus a dynamic inverter for right-inversion of a time-varying matrix is

$$\dot{\Gamma} = -\mu\Gamma \cdot (A(t)\Gamma - I) - \Gamma \dot{A}(t)\Gamma \tag{3.18}$$

Alternatively we may use Theorem 3.1 to invert $A(t)A(t)^T$, constructing the right inverse as $A(t)^T\Gamma(t)$.

In the case that $m \geq n$, $A(t)$ has a left inverse $\Gamma_*(t)$ which satisfies

$$F(\Gamma, t) := \Gamma A(t) - I = 0 \tag{3.19}$$

We may use the dynamic inverter (3.18) with $A(t)$ replaced by $A(t)^T$ to approximate the left inverse of $A(t)$.

4 Inversion of Constant Matrices

In this section we consider two methods for the dynamic inversion of constant matrices. In Section 6, relying on the methods of Section 5, we will consider another approach to the same problem.

Constant matrices may be inverted in a manner similar to the inversion of time-varying matrices as described in the last section. Let

$$F(\Gamma) := M\Gamma - I \tag{4.1}$$

Let $\Gamma(t)$ denote the estimator for the inverse of a constant matrix M, with $\Gamma_* = M^{-1}$ as the solution of $F(\Gamma) = 0$. Since M is constant, $\dot{\Gamma}_*$ is zero. As a consequence, if $\Gamma(0)$ is sufficiently close to Γ_*, then we can let $G[w, \Gamma] := \Gamma \cdot w$ and use the dynamic inverter

Dynamic Inverter for Constant Square Matrices

$$\dot{\Gamma} = -\mu\Gamma(M\Gamma - I) \tag{4.2}$$

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

4.1 A Comment on Gradient Methods

As shown in Section 3, Example 3.2, the function $G[w,\Gamma] := \Gamma \cdot w$ 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

Dynamic Inverter for Constant Square Matrices

$$\begin{gathered} \dot{\Gamma} = -\mu M^T(M\Gamma - I) \\ \Gamma \to M^{-1} \end{gathered} \tag{4.3}$$

Remark 4.1 Left and Right Inverses of Constant Matrices If M has full row-rank, with $M \in \mathbb{R}^{m\times n}$, $m \leq n$, then the equilibrium solution Γ_* of (4.3) is the right inverse $M^R := M^T(MM^T)^{-1}$ of M.

Dynamic Right-Inverter for Constant Matrices

$$\begin{gathered} \dot{\Gamma} = -\mu M^T(M\Gamma - I) \\ \Gamma \to M^L \end{gathered} \tag{4.4}$$

To obtain the right inverse of $M \in \mathbb{R}^{m\times n}$, $m \geq n$, where M has full column-rank, replace M in (4.4) by M^T. ▲

The dynamic inverter (4.3) is the standard least squares gradient flow (see [15], Section 1.6) for the function $\Phi : \mathbb{R}^n \to \mathbb{R}; \Gamma \mapsto \Phi(\Gamma)$ where

$$\Phi(\Gamma) := \frac{1}{2}\|M\Gamma - I\|_2^2 \tag{4.5}$$

It is also the neural-network constant matrix inverter of Wang [14]. Of course other gradient schemes may have the same solution as (4.3) though they may start from gradients of functions other than (4.5) (See, for instance [13]). In general, artificial neural networks are constructed to dynamically solve for the minimum of an energy function having a unique (at least locally) minimum, i.e. they realize gradient flows.

4.1.1 Connecting Gradient Methods with Dynamic Inversion

In general a dynamic inverter consists of three functions, F, G, and E as described in [1, 2]. The function $F(\Gamma, t)$ is the implicit function to be inverted, $G[w, \theta, t]$ is a dynamic inverse for $F(\Gamma, t)$, and $E(\theta, t)$ is an estimator for the derivative with respect to t of the root Γ_* of $F(\Gamma, t) = 0$. In order to relate gradient methods to dynamic inversion we consider the decomposition of a gradient flow system into an E, F, and G forming a dynamic inverter. For instance, let $H : \mathbb{R}^{n\times n} \times \mathbb{R} \to \mathbb{R}$ be a smooth function. A gradient system based on this function is

Gradient System

$$\dot{\Gamma} = -\nabla H(\Gamma, t) + \frac{\partial}{\partial t} H(\Gamma, t) \tag{4.6}$$

where ∇ denotes the gradient of $H(\Gamma, t)$. We may always identify gradient systems with dynamic inversion through the trivial dynamic inverse

$$G[w] = w \tag{4.7}$$

Then

$$F(\Gamma, t) = \nabla H(\Gamma, t) \tag{4.8}$$

and

$$E(\Gamma, t) = \frac{\partial}{\partial t} H(\Gamma, t) \tag{4.9}$$

Let $\mu = 1$. Then

$$\dot{\Gamma} = -G[F(\Gamma, t)] + E(\Gamma, t) \tag{4.10}$$

is the same as (4.6). Thus we have decomposed the gradient system (4.6) into an E, F, and G. It is more interesting, however, to find a dynamic inverse G such that if G were changed to the identity map, then the desired root would still be the solution to $F(\Gamma, t) = 0$, but the resulting dynamic inverter would not converge to the desired root. For example, identifying $F(\Gamma) = M\Gamma - I$, $G[w] = M^T \cdot w$, and $E = 0$ decomposes the gradient flow (4.3) into a dynamic inverter. For arbitrary $M \in GL(n, \mathbb{R})$ the stability properties of $\dot{\Gamma} = -\mu F(\Gamma)$ are unknown. But with G defined as $G[w] = M^T \cdot w$, $\dot{\Gamma} = -\mu G[F(\Gamma)]$ has an asymptotically stable equilibrium at $\Gamma_* = M^{-1}$. For a system of the form (4.3) such a decomposition is straightforward. For more complicated gradient systems however, we have no general methodology for decomposition into E, F, and G.

4.2 Dynamic Inversion of Constant Matrices by a Prescribed Time

The constant matrix dynamic inverters (4.2) and (4.3) above have the potential disadvantage of producing an exact inverse only asymptotically as $t \to \infty$. One may, however, wish to obtain the inverse by a prescribed time. To obtain inversion by a prescribed time 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 finite time $t > 0$, say $t = 1$, then perhaps we could use the technique of Section 3 for the inversion of time-varying matrices in order 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 4.2 Constant Matrix Inversion by a Prescribed Time Using Homotopy. Let M be a constant matrix in $\mathbb{R}^{n\times n}$. We wish to dynamically determine the inverse of M. Consider the t-dependent matrix,

Matrix Homotopy

$$H(t) = (1 - t)I + tM. \tag{4.11}$$

In the space of $n \times n$ matrices, $t \mapsto H(t)$ describes a parameterized curve, or homotopy, of matrices from the identity to $M = H(1)$ as indicated in Figure 1; in fact this curve (4.11) is a straight line.

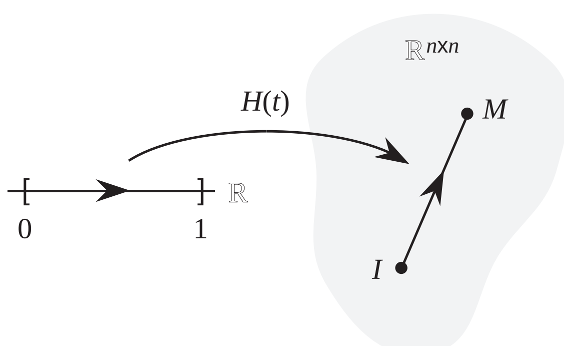


Figure 1: The matrix homotopy $H(t)$.

From Theorem 3.1 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 substituting $H(t)$ for $A(t)$, and $\dot{H}(t) = M - I$ for $\dot{A}(t)$ in (3.8). Since our initial conditions are a precise inverse of $H(0)$, Theorem 3.1 tells us that the matrix Γ becomes the precise inverse of M at time $t = 1$ as shown schematically in Figure 2. *That is, of course, if $H(t)$ remains nonsingular as t goes from 0 to 1!*

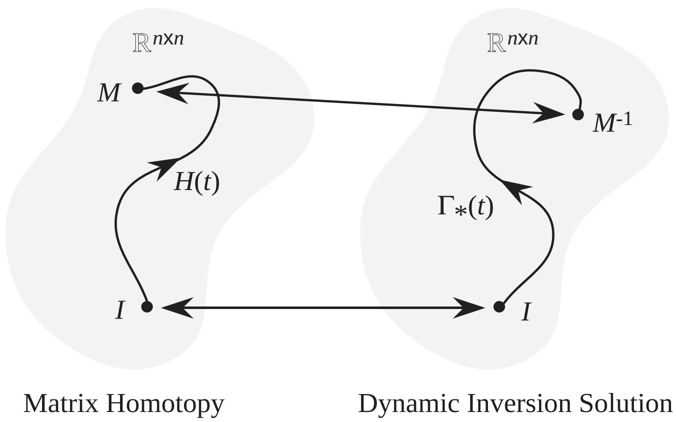


Figure 2: The matrix homotopy $H(t)$ from I to M with the corresponding solution $\Gamma_*(t)$, the inverse of $H(t)$.

▲

For a dynamic inverter for this example let

$$\begin{aligned} F(\Gamma, t) &:= ((1-t)I + tM)\,\Gamma - I \\ G[w, \Gamma] &:= \Gamma \cdot w \\ E(\Gamma) &:= -\Gamma(M - I)\Gamma \end{aligned} \tag{4.12}$$

Then a dynamic inverter is $\dot{\Gamma} = -\mu G[F(\Gamma, t), \Gamma] + E(\Gamma)$ with $\Gamma(0) = I$. Expanded, this is

Prescribed-Time Dynamic Inverter for Constant Matrices

$$\dot{\Gamma} = -\mu\Gamma\left(((1-t)I + tM)\Gamma - I\right) - \Gamma(M - I)\Gamma$$

Another choice of linear dynamic inverse is $G[w,t] := ((1-t)I + tM)^T$ giving

Prescribed-Time Dynamic Inverter for Constant Matrices

$$\dot{\Gamma} = -\mu H(t)^T\,(H(t)\Gamma - I) - \Gamma(M - I)\Gamma$$

▲

Homotopy-based methods, also called continuation methods, for solving sets of linear and nonlinear equations have been around for quite some time. For a review of developments prior to 1980 see Allgower and Georg [16] The general idea is that one starts with a problem with a known solution (e.g. the inverse of the identity matrix) and smoothly transforms that problem to a problem with an unknown solution, transforming the known solution in a corresponding manner until the unknown solution is reached. Often it is considerably easier to transform a known solution to a problem into an unknown solution to a closely related problem rather than calculating the new solution from scratch. Solution of the roots of nonlinear polynomial equations (see Dunyak et al. [17] and Watson [18] for examples) is a typical example with broad engineering application.

Remark 4.3 Requirement for Nonsingular Homotopy. The scheme of Example 4.2 requires that there is no $t \in [0,1]$ for which $H(t)$ (4.11) is singular.

▲

Recall that there are two maximal connected open subsets which comprise $GL(n,\mathbb{R})$, namely $GL^+(n,\mathbb{R}) = \{M \in \mathbb{R}^{n\times n} | \det(M) > 0\}$ and $GL^-(n,\mathbb{R}) = \{M \in \mathbb{R}^{n\times n} | \det(M) < 0\}$. These two sets are disjoint and are separated by the codimension-1 manifold of singular $n \times n$ matrices $\{M \in \mathbb{R}^{n\times n} | \det(M) = 0\}$. The identity I is in $GL^+(n,\mathbb{R})$. In order for the curve $t \mapsto H(t)$ to be dynamically invertible, it must never leave $GL^+(n,\mathbb{R})$ (see Figure 3).

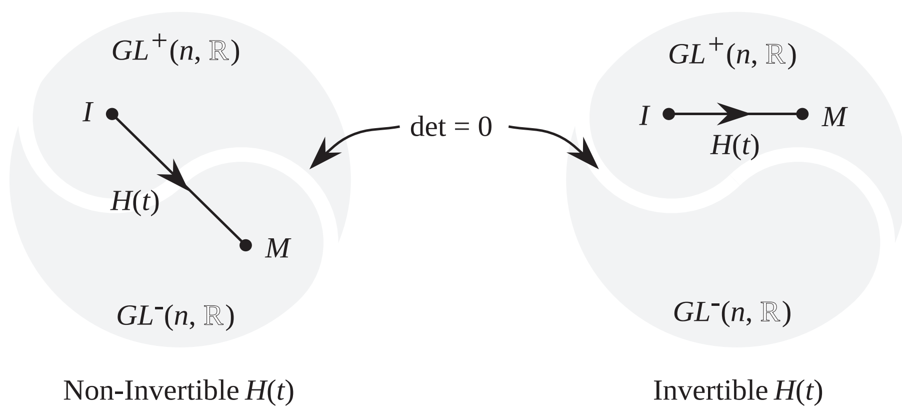


Figure 3: The homotopy from I to M must remain in $GL^+(n,\mathbb{R})$ to be invertible.

For our particular choice of $H(t)$, since $H(0)=I$, and I is in $GL^+(n,\mathbb{R})$, the homotopy $H(t)$ must be confined to $GL^+(n,\mathbb{R})$ to be invertible for all $t\in[0,1]$. The following lemma specifies sufficient conditions on M for $H(t)$ (4.11) to remain in $GL^+(n,\mathbb{R})$ as t goes from 0 to 1.

Lemma 4.4 Matrix Homotopy Lemma. *If $M\in GL(n,\mathbb{R})$ has no eigenvalues in $(-\infty,0)$, then for each $t\in[0,1]$, $H(t)=(1-t)I+tM$ is in $GL(n,\mathbb{R})$.* ♦

Remark 4.5 Inversion of Positive-Definite Symmetric Constant Matrices. If M is a positive-definite symmetric matrix, then the assumption of Lemma 4.4 holds. ▲

Proof of Lemma 4.4: Suppose that $H(t)=(1-t)I+tM$ is singular for some $\bar{t}\in[0,1]$. The identity I is nonsingular as is M by assumption, so $\bar{t}\notin\{0,1\}$. Thus there exists a non-zero $v\in\mathbb{R}^n$ such that

$$((1-\bar{t})I+\bar{t}M)\,v=0 \tag{4.13}$$

Since $\bar{t}\neq 0$ we can divide (4.13) by $-\bar{t}$ to obtain

$$\left(\frac{(\bar{t}-1)}{\bar{t}}I-M\right)v=0 \tag{4.14}$$

But $\bar{t}$ can only satisfy (4.14) if $\lambda(\bar{t}):=(\bar{t}-1)/\bar{t}$ is an eigenvalue of M. As t ranges over $(0,1)$, $\lambda(t)$ ranges over $(-\infty,0)$. But by assumption M has no eigenvalues in $(-\infty,0)$, hence no such $\bar{t}$ exists in $(0,1)$ and $H(t)$ is nonsingular on $[0,1]$. ■

We may obtain the exact inverse of M at *any* prescribed time $t_1 > 0$ by a slight modification of the homotopy (4.11). We summarize our results of this section in the following theorem.

Theorem 4.6 Dynamic Inversion of Constant Matrices by a Prescribed Time. *For any constant $M \in GL(n, \mathbb{R})$, and for any prescribed $t_1 > 0$, if $\sigma(M) \cap (-\infty, 0] = \varnothing$, then the solution $\Gamma(t)$ of the dynamic inverter*

Prescribed-Time Dynamic Inverter for Constant Matrices

$$\dot{\Gamma} = -\mu\Gamma\left(\left((1-\frac{t}{t_1})I + \frac{t}{t_1}M\right)\Gamma - I\right) - \Gamma(M-I)\Gamma \tag{4.15}$$

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

Remark 4.7 Preservation of Symmetry. If M is symmetric, then the right-hand side of (4.15) is also symmetric. Thus if $\Gamma(0)$ is symmetric, then $\Gamma(t)$ is symmetric for all t. ▲

Example 4.8 Right and Left Inverses of Constant Matrices by a Prescribed Time. Let $A \in \mathbb{R}^{m\times n}$ be a constant matrix with $m \leq n$ and assume that A has full row rank. The right inverse of A is given by $A^R := A^T(AA^T)^{-1}$. To obtain the right inverse A^R at time t_1, we may apply Theorem 4.6 replacing M by AA^T which is positive definite. Then $A^T(AA^T)^{-1} = A^T\Gamma(t_1)$.

Prescribed-Time Dynamic Right-Inversion of a Constant Matrix

$$\dot{\Gamma} = -\mu\Gamma\left(\left((1-\tfrac{t}{t_1})I + \tfrac{t}{t_1}AA^T\right)\Gamma - I\right) - \Gamma(AA^T - I)$$
$$A^T\Gamma(t_1) = A^R$$

When a constant A has full column rank, then since A^TA is positive definite, the left inverse $A^R := (A^TA)^{-1}A^T$ may be obtained by substituting A^TA for

M in Theorem 4.6. Then $A^L = \Gamma(t_1)A^T$.

Prescribed-Time Dynamic Left-Inversion of a Constant Matrix

$$\dot{\Gamma} = -\mu\Gamma\left(\left((1-\tfrac{t}{t_1})I + \tfrac{t}{t_1}A^TA\right)\Gamma - I\right) - \Gamma(A^TA - I)$$
$$\Gamma(t_1)A^T = A^L$$

▲

Theorem 4.6 is limited in its utility by the necessity that M have a spectrum which does not intersect $(-\infty, 0)$. By appealing to the polar decomposition in Section 6 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 constant M, irrespective of its spectrum, by any prescribed time $t_1 \geq 0$.

5 Polar Decomposition for Time-Varying Matrices

In this section we will show how dynamic inversion may be used to perform polar decomposition [19] 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 on $(0, \infty)$.

Though polar decomposition will be used here largely as a path to inversion, polar decomposition finds substantial utility in its own right. In particular it is used widely in the study of stress and strain in continuous media. See, for instance, Marsden and Hughes [20].

First consider the polar decomposition of a *constant* matrix $M \in GL(n, \mathbb{R})$, $M = PU$ where U is in the space of $n \times n$ orthogonal matrices with real entries, $O(n, \mathbb{R})$, 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.

Now consider the case of a t-dependent nonsingular square matrix $A(t)$. Since $A(t)$ is nonsingular for all t, $A(t)A(t)^T$ is positive-definite. For any t, the unique positive-definite solution to $XA(t)A(t)^TX - I = 0$ is $X_*(t) = P(t)^{-1}$. If we have $X_*(t) = P(t)^{-1}$, then from $A(t) = P(t)U(t)$ we can get the orthogonal factor $U(t)$ of the polar decomposition by $U(t) = X_*(t)A(t)$, as well as the symmetric part $P(t)$ of the polar decomposition by $P(t) = X_*(t)A(t)A(t)^T$. We can also obtain the inverse of $A(t)$ as $A(t)^{-1} = U(t)^TX_*(t)$.

Since $P(t)$ is a symmetric $n \times n$ matrix, it is parameterized by

$$s(n) := n(n+1)/2 \tag{5.1}$$

elements as is its inverse $P^{-1}(t)$. We will construct the dynamic inverter that produces $P^{-1}(t)$.

Remark 5.1 Vector Notation for Symmetric Matrices. It will be convenient for the purposes of this section and the next to adopt a notation that allows us to switch between matrix representation and vector representation of elements of $S(n,\mathbb{R})$. The convenience of this notation will be seen in Section 5.1 to arise from the lack of a convenient matrix form of the inverse of the linear matrix mapping on $S(n,\mathbb{R})$, $X \mapsto XM + MX$ where X and M are in $S(n,\mathbb{R})$.

Choose an ordered basis

$$\beta = \{\beta_1, \ldots, \beta_{s(n)}\} \tag{5.2}$$

for $S(n,\mathbb{R})$. For any $x \in \mathbb{R}^{s(n)}$ there corresponds a unique matrix $\hat{x} \in S(n,\mathbb{R})$ where the correspondence is through the expansion of $\hat{x}$ in the ordered basis β,

$$\hat{x} \equiv (x)\hat{} := \sum_{i \in s(n)} x^i \beta_i \in S(n,\mathbb{R}) \tag{5.3}$$

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

$$X = \sum_{i \in \underline{s(n)}} x^i \beta_i \tag{5.4}$$

in the basis β so that

$$\check{X} \equiv (X)\check{} = x \tag{5.5}$$

Then

$$(\check{X})\hat{} = X \quad \text{and} \quad (\hat{x})\check{} = x \tag{5.6}$$

▲

Let

$$\Lambda(t) := A(t)A(t)^T \tag{5.7}$$

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

$$F(x,t) := (\hat{x}\Lambda(t)\hat{x} - I)\check{} \tag{5.8}$$

Let x_* be a solution of $F(x,t) = 0$. Then $\hat{x}_*$ is a symmetric square root of $\Lambda(t)$.

Nothing in the form of $F(x,t)$ (5.8) enforces the *positive-definiteness* of the solution $\hat{x}_*(t)$, where $x_*(t)$ is the solution of $F(x,t) = 0$. For instance, for each solution $x_*(t)$ of $F(x,t) = 0$, $-x_*(t)$ is also a solution. Each solution $t \mapsto x_*(t)$ is, however, *isolated* as long as $D_1F(x_*,t)$, where $F(x,t)$ is defined by (5.8), is nonsingular. We will show in the next subsection, Subsection 5.1, that the nonsingularity of $A(t)$ implies the nonsingularity of $D_1F(x_*,t)$.

5.1 The Lyapunov Map

We will use a linear dynamic inverse for $F(x,t)$ (5.8) 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 $D_1F(x_*,t)$.

Differentiate

$$\check{F}(x,t) = \check{x}\Lambda(t)\check{x} - I \tag{5.9}$$

with respect to $\hat{x}$ to get

$$(D_1F(x,t))\check{} = \frac{1}{2}\left(\hat{x}\Gamma(t) + \Gamma(t)\hat{x}\right) \tag{5.10}$$

The differential $D_1F(x,t)$ expressed as a mapping $S(n,\mathbb{R}) \to S(n,\mathbb{R})$ is

$$L_{\Lambda(t)\hat{x}} : Y \mapsto L_{\Lambda(t)\hat{x}}(Y) := \frac{1}{2}\left(Y\Lambda(t)\hat{x} + \hat{x}\Lambda(t)Y\right) \tag{5.11}$$

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

$$L_M : Y \mapsto L_M Y := YM + MY \tag{5.12}$$

with Y and M in $\mathbb{R}^{n\times n}$ as a *Lyapunov map* due to its relation to the *Lyapunov equation* $YM + MY = Q$ which arises in the study of the stability of linear control systems (see e.g. Horn and Johnson [21], Chapter 4). It may be easily verified that a Lyapunov map (5.12) is linear in Y. It may also be proven that L_M is an invertible map if no two eigenvalues of M add up to zero (see e.g. [21], Theorem 4.4.6, page 270).

Now note that $\Lambda(t)\hat{x}_* = \hat{x}_*\Lambda(t) = P(t)$ which is positive-definite and symmetric. Thus no pair of eigenvectors of $\Lambda(t)\hat{x}_*$ sum to zero. Therefore $L_{\Lambda(t)\hat{x}_*}(Y)$ is nonsingular. It follows then that the matrix $D_1F(x_*,t)$ is invertible. Since $D_1F(x,t)$ is continuous in x, it follows that $D_1F(x,t)$ remains invertible for all x in a sufficiently small neighborhood of x_*.

Though numerical inversion of the Lyapunov map has long been a topic of interest in the context of control theory [22, 23], we do not know of any *matrix* map $L^{-1} : S(n,\mathbb{R}) \to S(n,\mathbb{R})$, taking *matrices* to *matrices*, which inverts L_M. By converting L_M to an $s(n)\times s(n)$ matrix, however, and representing elements of $S(n,\mathbb{R})$ as vectors, the inverse L^{-1} as a mapping between vector spaces $\mathbb{R}^{s(n)} \to \mathbb{R}^{s(n)}$ can be obtained through standard matrix inversion or, as we will see, dynamic matrix inversion. This is why we will sometimes resort to the vector notation of Remark 5.1 in referring to elements of $S(n,\mathbb{R})$.

5.2 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

$$D_1F(x_*,t)^{-1} = \Gamma_* \tag{5.13}$$

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

$$G[w,\Gamma] := D_1F(x_*,t)^{-1}\big|_{\Gamma_*=\Gamma} \cdot w = \Gamma \cdot w \tag{5.14}$$

This makes $G[w,\Gamma]$ (5.14) a dynamic inverse for $F(x,t) = (\hat{x}\Lambda(t)\hat{x} - I)\check{}$.

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

$$D_1F(x_*,t)\dot{x}_* + D_2F(x_*,t) = 0 \tag{5.15}$$

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

$$\dot{x}_* = -D_1F(x_*,t)^{-1}D_2F(x_*,t) = -\Gamma_* D_2F(x_*,t) \tag{5.16}$$

Note that $D_2F(x_*,t) = (\hat{x}_*\dot{\Lambda}(t)\hat{x}_*)\check{}$. Now, substituting x and Γ for x_* and Γ_* we obtain

$$E(x,\Gamma,t) := -\Gamma\left(\hat{x}\dot{\Lambda}(t)\hat{x}\right)\check{} \tag{5.17}$$

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

$$F^\gamma(x,\Gamma,t) := D_1F(x,t)\Gamma - I \tag{5.18}$$

A linear dynamic inverse for $F^\gamma(x,\Gamma,t)$ is $G^\gamma : \mathbb{R}^{s(n)\times s(n)} \times \mathbb{R}^{s(n)\times s(n)} \to \mathbb{R}^{s(n)\times s(n)}$; $(w,\Gamma) \mapsto G^\gamma[w,\Gamma]$ defined by

$$G^\gamma[w,\Gamma] := \Gamma \cdot w \tag{5.19}$$

For an estimator $E^\gamma(x,\Gamma,t)$ for $\dot{\Gamma}_*$, we differentiate $F^\gamma(x_*,\Gamma_*,t) = 0$ with respect to t, solve for $\dot{\Gamma}_*$, and substitute x and Γ for x_* and Γ_* respectively to get

$$E^\gamma(x,\Gamma,t) := -\Gamma\left(\frac{d}{dt}D_1F(x,t)\right)\Bigg|_{\dot{x}_*=E(x,\Gamma,t)} \cdot \Gamma \tag{5.20}$$

Combining the E's, F's, and G's from (5.17), (5.8), (5.14), (5.20), (5.18), and (5.19), we obtain the dynamic inverter

$$\left\{\begin{array}{rcl} \dot{x} & = & -\mu G[F(x,t),\Gamma] + E(\Gamma,x,t) \\ \dot{\Gamma} & = & -\mu G^\gamma[F^\gamma(x,\Gamma,t),\Gamma] + E^\gamma(x,\Gamma,t) \end{array}\right. \tag{5.21}$$

or in an expanded form

Dynamic Polar Decomposition for Time-Varying Matrices

$$\left\{\begin{array}{rcl} \dot{x} & = & -\mu\Gamma\left(\hat{x}\Lambda(t)\hat{x}-I\right)^{\check{}} - \Gamma\left(\hat{x}\dot{\Lambda}(t)\hat{x}\right)^{\check{}} \\ \dot{\Gamma} & = & -\mu\Gamma\left(D_1F(x,t)\Gamma - I\right) \\ & & -\Gamma\left(\frac{d}{dt}D_1F(x,t)\right)\Big|_{\dot{x}_*=E(x,\Gamma,t)}\cdot\Gamma \end{array}\right.$$

$$\begin{gathered} \hat{x}A(t) \to U(t) \\ \hat{x}A(t)A(t)^T \to P(t) \\ A(t)^T(\hat{x})^2 \to A^{-1}(t) \end{gathered} \tag{5.22}$$

Initial conditions for the dynamic inverter may be set so that $\hat{x}(0) \approx P(0)^{-1}$ and $\Gamma(0) \approx D_1F((P(0)^{-1})^{\check{}}, t)^{-1}$.

Combining the results above with the dynamic inversion theorem, Theorem 2.3 gives the following theorem.

Theorem 5.2 Dynamic Polar Decomposition of Time-Varying Matrices. *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 (5.22) where $F(x,t)$ is given by (5.8). Then there exists a $\tilde{\mu}$ such that if $\mu > \tilde{\mu}$, and $(\hat{x}(0),\Gamma(0))$ is sufficiently close to $(P(0)^{-1}, D_1F((P(0)^{-1})^{\check{}},t)^{-1})$, then*

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

◆

An example of the polar decomposition of a 2×2 matrix will illustrate application of Theorem 5.2.

Example 5.3 Polar Decomposition of a Time-Varying Matrix. Let

$$A(t) := \begin{bmatrix} 10+\sin(10t) & \cos(t) \\ -t & 1 \end{bmatrix} \tag{5.23}$$

In this case $x \in \mathbb{R}^3$ and $\Gamma \in \mathbb{R}^{3\times 3}$. We will perform polar decomposition and inversion of $A(t)$ over $t \in [0,8]$, an interval over which $A(t)$ is nonsingular.

We will estimate $P(t)$ and $U(t)$ such that $A(t) = P(t)U(t)$, with $P(t) \in S(2,\mathbb{R})$ being the positive-definite symmetric square root of $A(t)A(t)^T$, and with $U(t) \in O(2,\mathbb{R})$.

We choose the ordered basis β of $S(2,\mathbb{R})$ to be

$$\beta = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\} \tag{5.24}$$

In this basis we have

$$F(x,t) = \begin{bmatrix} \lambda_1 x_1^2 + 2\lambda_2 x_1 x_2 + \lambda_3 x_2^2 - 1 \\ \lambda_1 x_1 x_2 + \lambda_2 x_2^2 + \lambda_2 x_1 x_3 + \lambda_3 x_2 x_3 \\ \lambda_1 x_2^2 + 2\lambda_2 x_2 x_3 + \lambda_3 x_3^2 - 1 \end{bmatrix} \tag{5.25}$$

Then

$$D_1 F(x,t) = \\ \frac{1}{2} \begin{bmatrix} 2(\lambda_1 x_1 + \lambda_2 x_2) & 2(\lambda_2 x_1 + \lambda_3 x_2) & 0 \\ \lambda_1 x_2 + \lambda_2 x_3 & \lambda_1 x_1 + 2\lambda_2 x_2 + \lambda_3 x_3 & \lambda_2 x_1 + \lambda_3 x_2 \\ 0 & 2(\lambda_1 x_2 + \lambda_2 x_3) & 2(\lambda_2 x_2 + \lambda_3 x_3) \end{bmatrix} \tag{5.26}$$

$$E(x,\Gamma,t) = -\frac{1}{2}\Gamma \begin{bmatrix} \dot{\lambda}_1 x_1^2 + 2\dot{\lambda}_2 x_1 x_2 + \dot{\lambda}_3 x_2^2 \\ \dot{\lambda}_1 x_1 x_2 + \dot{\lambda}_2 x_2^2 + \dot{\lambda}_2 x_1 x_3 + \dot{\lambda}_3 x_2 x_3 \\ \dot{\lambda}_1 x_2^2 + 2\dot{\lambda}_2 x_2 x_3 + \dot{\lambda}_3 x_3^2 \end{bmatrix} \tag{5.27}$$

and

$$\left. \frac{d}{dt} D_1 F(x_*, t) \right|_{x_* = x,\ \dot{x}_* = E(x,\Gamma,t)} = \frac{1}{2} \begin{bmatrix} \dot{L}_{11} & \dot{L}_{12} & 0 \\ \dot{L}_{21} & \dot{L}_{22} & \dot{L}_{23} \\ 0 & \dot{L}_{32} & \dot{L}_{33} \end{bmatrix} \tag{5.28}$$

where

$$\begin{aligned} \dot{L}_{11} &= 2\dot{\lambda}_1 x_1 + 2\lambda_1 E_1(x,\Gamma,t) + 2\dot{\lambda}_2 x_2 + 2\lambda_2 E_2(x,\Gamma,t) \\ \dot{L}_{12} &= 2\dot{L}_{23} \\ \dot{L}_{21} &= \dot{\lambda}_1 x_2 + \lambda_1 E_2(x,\Gamma,t) + \dot{\lambda}_2 x_3 + \lambda_2 E_3(x,\Gamma,t) \\ \dot{L}_{22} &= \dot{\lambda}_1 x_1 + \lambda_1 E_1(x,\Gamma,t) + 2\dot{\lambda}_2 x_2 + 2\lambda_2 E_2(x,\Gamma,t) \\ &\quad + \dot{\lambda}_3 x_3 + \lambda_3 E_3(x,\Gamma,t) \\ \dot{L}_{23} &= \dot{\lambda}_2 x_1 + \lambda_2 E_1(x,\Gamma,t) + \dot{\lambda}_3 x_2 + \lambda_3 E_2(x,\Gamma,t) \\ \dot{L}_{32} &= 2\dot{L}_{21} \\ \dot{L}_{33} &= 2\dot{\lambda}_2 x_2 + 2\lambda_2 E_2(x,\Gamma,t) + 2\dot{\lambda}_3 x_3 + 2\lambda_3 E_3(x,\Gamma,t) \end{aligned} \tag{5.29}$$

Dynamic inversion using Equations (5.22) was simulated using the adaptive step size Runge-Kutta integrator `ode45` from Matlab, with the default tolerance of 10^{-6}. The initial conditions were set so that

$$\begin{aligned} \hat{x}(0) &= \Lambda^{1/2}(0) + \hat{e}_x \\ \Gamma(0) &= D_1 F(x(0),t)^{-1} \end{aligned} \tag{5.30}$$

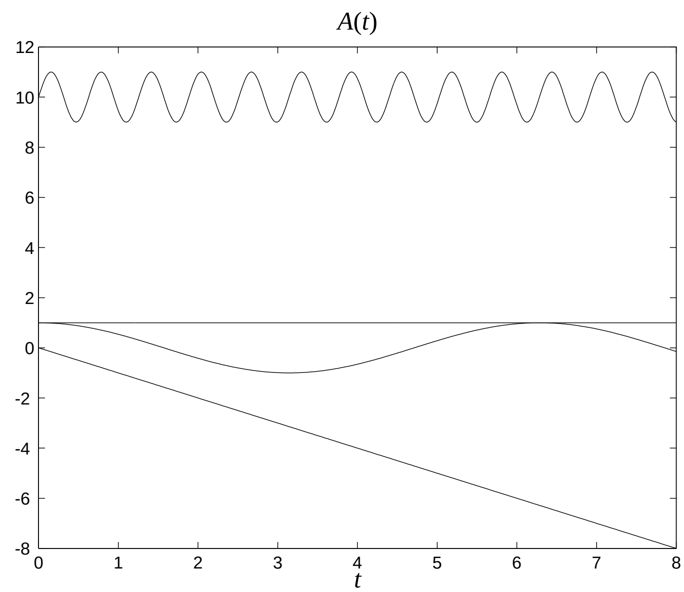


Figure 4: Elements of $A(t)$ (see (5.23)). See Example 5.3

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.

The graph of Figure 4 shows the values of the individual elements of $A(t)$. The top graph of Figure 5 shows the elements of $x(t)$, the estimator for $P(t)^{-1}$, and the bottom graph of Figure 5 shows the elements of $\Gamma(t)$.

Figure 6 shows $\log_{10}(\|\hat{x}(t)\Lambda(t)\hat{x}(t) - I\|_\infty)$ indicating the extent to which $\hat{x}$, the estimator for $P(t)^{-1}$ fails to be the square root of $\Lambda(t) = A(t)A(t)^T$.

For estimates of $P(t)$, $U(t)$, and $A(t)^{-1}$ we have

$$\Gamma(t)A(t)A(t)^T \to P(t),\ \Gamma(t)A(t) \to U(t),\ \text{and } A(t)^T\Gamma^2 \to A(t)^{-1} \tag{5.31}$$

▲

Remark 5.4 Symmetry of the Dynamic Inverter. It is interesting to note that $P(t)^{-1}$, besides being a solution to $\hat{x}\Lambda(t)\hat{x} - I = 0$ is also a solution to $\Lambda(t)\hat{x}^2 - I = 0$ as well as $\hat{x}^2\Lambda(t) - I = 0$. But $\Lambda(t)\hat{x}^2 - I$ and $\hat{x}^2\Lambda(t) - I$ are not, in general, symmetric even when $\Lambda(t)$ and $\hat{x}$ are symmetric. Though exponential convergence is still guaranteed when using these forms, the flow $\Gamma(t)$ is not, in general, confined to $S(n, \mathbb{R})$. Using these forms would increase the number of

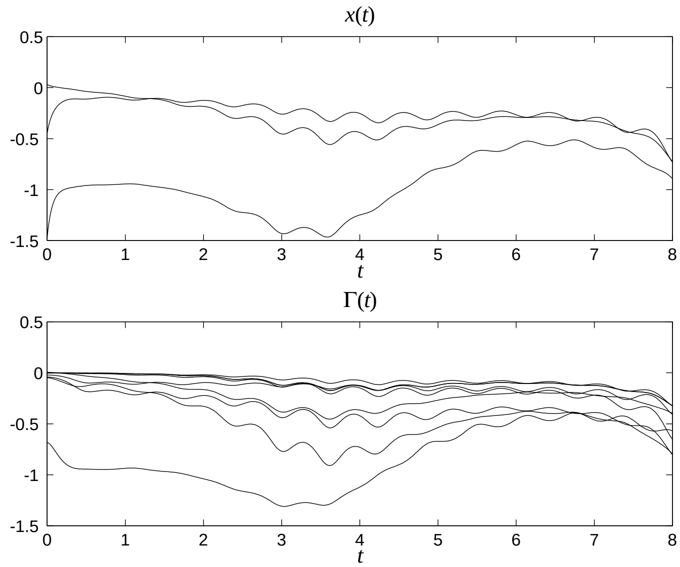


Figure 5: Elements of x (top), and Γ (bottom). See Example 5.3.

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 (5.22) no longer be symmetric, but Γ would be $n^2 \times n^2$ rather than $s(n) \times s(n)$. ▲

6 Polar Decomposition and Inversion of Constant Matrices

In the dynamic inversion techniques of Sections 3 and 5 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 constant matrix, $A(0)$, in order to start the dynamic inverter. Methods of constant matrix inversion presented in Section 4 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 the interval $(-\infty, 0)$. The question naturally arises then, how might we use dynamic inversion to invert *any* constant matrix so that the exact inverse is available by a prescribed 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

$$P = P^T > 0, \quad UU^T = I, \quad \text{and} \quad M = PU \tag{6.1}$$

Helmke and Moore (see [15], pages 150-152) have described a gradient flow for

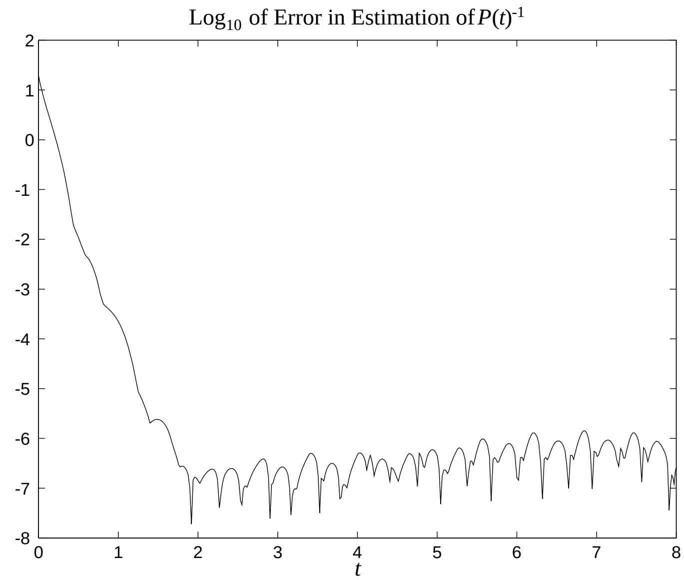


Figure 6: The error $\log_{10}(\|\hat{x}(t)\Lambda(t)\hat{x}(t) - I\|_\infty)$ indicating the extent to which x fails to satisfy $\hat{x}\Lambda(t)\hat{x} - I = 0$. The ripple from $t \approx 1.8$ to $t = 8$ is due to numerical noise. See Example 5.3.

the function $\|A - UP\|$,

$$\begin{aligned} \dot{\bar{U}} &= \bar{U}\bar{P}M^T\bar{U} - M\bar{P} \\ \dot{\bar{P}} &= -2\bar{P} + M^T\bar{U} + \bar{U}^T M \end{aligned} \tag{6.2}$$

where $\bar{P}$ and $\bar{U}$ are meant to approximate P and U respectively. Asymptotically, this system produces factors P_* and U_* satisfying $M - P_*U_* = 0$ for almost all initial conditions $\bar{P}(0), \bar{U}(0)$ as $t \to \infty$. A difficulty with this approach, as the authors point out, is that positive-definiteness of the approximator $\bar{P}$ is not guaranteed.

In this section we describe a dynamic system that provides polar decomposition of *any* nonsingular constant matrix by any prescribed time, with the positiveness of the estimator of P guaranteed. This will be accomplished by applying Theorem 5.2 on dynamic polar decomposition of time-varying matrices to the homotopy

$$\Lambda(t) := (1-t)I + tMM^T \tag{6.3}$$

Unlike the homotopy $H(t) = (1-t)I + tM$ of Section 4, the homotopy $\Lambda(t)$ (6.3) is guaranteed to have a spectrum which avoids $(-\infty, 0)$ for *any* nonsingular M since $\Lambda(t)$ is a positive definite symmetric matrix for all $t \in [0,1]$. The situation is depicted in Figure 7.

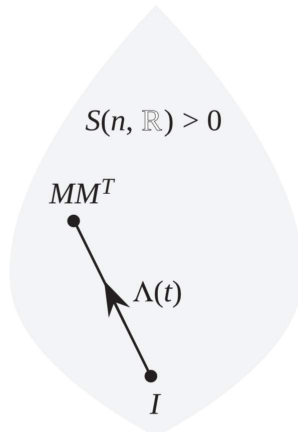


Figure 7: $H(t)$ is positive-definite and symmetric for all $t \in [0, 1]$.

Recall that M is in $GL(n, \mathbb{R})$. For $\Lambda(t)$ as defined in (6.3) note that $\Lambda(0) = I$, $\Lambda(1) = MM^T$, and for all $t \in [0, 1]$, $\Lambda(t)$ is positive-definite and symmetric. Let $P(t)$ denote the positive-definite symmetric square root of $\Lambda(t)$. Let the estimator of $P^{-1}(t)$ be $\hat{x} \in \mathbb{R}^{n \times n}$. Differentiate $\Lambda(t)$ (6.3) with respect to t to get

$$\dot{\Lambda}(t) = MM^T - I \tag{6.4}$$

Now we may apply the dynamic inverter of Section 5 in order to perform the polar decomposition of M. As in (5.8), let

$$F(x, t) := (\hat{x}\Lambda(t)\hat{x} - I)^{\vee} \tag{6.5}$$

By inspection it may be verified that $\hat{x}_*(0) = I$ and $\Gamma_*(0) = I$. If we set $\hat{x}(0) = I$ and $\Gamma(0) = I$, then Theorem 2.3 and the results of the last section assure us that $\hat{x}(t) \equiv P^{-1}$ for all $t \geq 0$, and thus $\hat{x}(1) = P^{-1}$. Consequently

$$\begin{array}{rcl} \hat{x}(1) & = & P^{-1} \\ \Lambda(1)\hat{x}(1) = MM^T\hat{x}(1) & = & P \\ \hat{x}(1)M & = & U \\ M^T\hat{x}(1)^2 & = & M^{-1} \end{array} \tag{6.6}$$

Note that $\dot{\Lambda}(t) = MM^T - I = 0$ 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 6.1 Dynamic Polar Decomposition of Constant Matrices by a Prescribed Time. *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 $x(0) = \check{I}$ and $\Gamma(0) = I$. Let $(x(t), \Gamma(t))$ denote the solution of*

Prescribed-Time Dynamic Inverter for Constant Matrices

$$\left\{ \begin{array}{rcl} \dot{x} & = & -\mu G[F(x,t), \Gamma] + E(x, \Gamma) \\ \dot{\Gamma} & = & -\mu G^{\gamma}[F^{\gamma}(\Gamma, x)] + E^{\gamma}(x, \Gamma). \end{array} \right.$$
$$\begin{gathered} \Lambda(t) = (1-t)I + tMM^T \\ F(x,t) = (\hat{x}\Lambda(t)\hat{x} - I)\check{} \\ F^{\gamma}(x, \Gamma, t) = D_1 F(x,t)\Gamma - I \\ G[w, \Gamma] = \Gamma \cdot w \\ G^{\gamma}[w, \Gamma] = \Gamma \cdot w \\ E(x, \Gamma) = -\Gamma(\hat{x}(MM^T - I)\hat{x})\check{} \\ E^{\gamma}(x, \Gamma) = -\Gamma \left(\tfrac{d}{dt} D_1 F(x,t)\right)\Big|_{\dot{x}=E(x,\Gamma)} \cdot \Gamma \end{gathered} \tag{6.7}$$

Then for any $\mu > 0$,

$$MM^T\hat{x}(1) = P, \quad \hat{x}(1)M = U, \quad \text{and} \quad M^T\hat{x}(1)^2 = M^{-1} \tag{6.8}$$

▲

Remark 6.2 Polar Decomposition by Any Prescribed Time. As in Theorem 4.6 we can force $\hat{x}$ 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 $\hat{x}(t_1) = P^{-1}$. ▲

Example 6.3 A digital computer simulation of a dynamic inverter for the polar decomposition of a constant 2-by-2 matrix was performed. The integration was performed in Matlab [24] using `ode45` an adaptive step size Runge-Kutta routine using the default tolerance of 10^{-6}. The matrix M was chosen randomly to be

$$M = \begin{bmatrix} 7 & -3 \\ -24 & -3 \end{bmatrix} \tag{6.9}$$

The value of μ was set to 10. The evolution of the elements of $x(t)$ and $\Gamma(t)$ are shown in Figure 8.

Figure 9 shows the base 10 log of $\|\hat{x}(t)MM^T\hat{x}(t) - I\|_\infty$ indicating the extent to which x, the estimator for P^{-1} fails to be the square root of $\Lambda(t) = MM^T$.

Final values ($t = 1$) of the error $\|\hat{x}(t)MM^T\hat{x}(t) - I\|_\infty$ was

$$\|\hat{x}(1)\Lambda(1)\hat{x}(1) - I\|_\infty = 1.0611 \times 10^{-6} \tag{6.10}$$

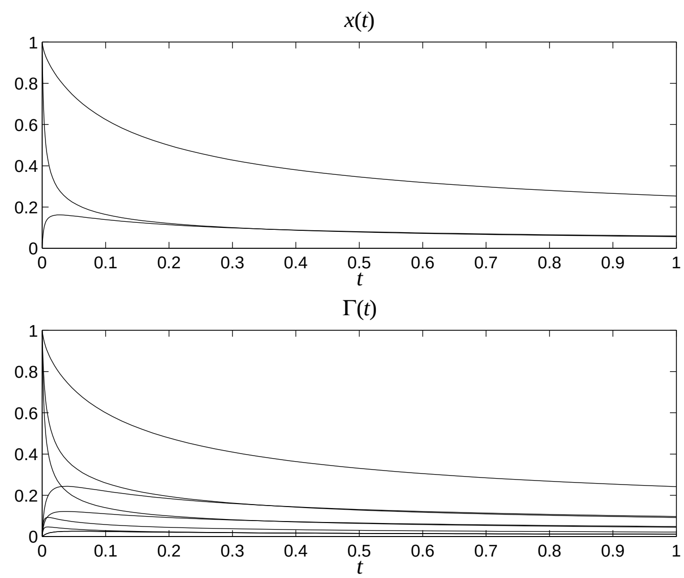


Figure 8: Elements of $x(t)$ (top) and $\Gamma(t)$ (bottom), for Example 6.3.

Final values of P, U, and A^{-1} were

$$\begin{aligned} P = MM^T\hat{x}(1) &= \begin{bmatrix} 5.2444 & -5.5223 \\ -5.5223 & 23.5479 \end{bmatrix} \\ U = \hat{x}(1)M &= \begin{bmatrix} 0.3473 & -0.9377 \\ -0.9377 & -0.3473 \end{bmatrix} \\ M^{-1} = M^T\hat{x}(1)^2 &= \begin{bmatrix} 0.0323 & -0.0323 \\ -0.2581 & -0.0753 \end{bmatrix} \end{aligned} \tag{6.11}$$

▲

7 Chapter Summary

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

In some control problems, dynamic inversion may provide essential signals which can be incorporated into controllers for nonlinear dynamic systems [25].

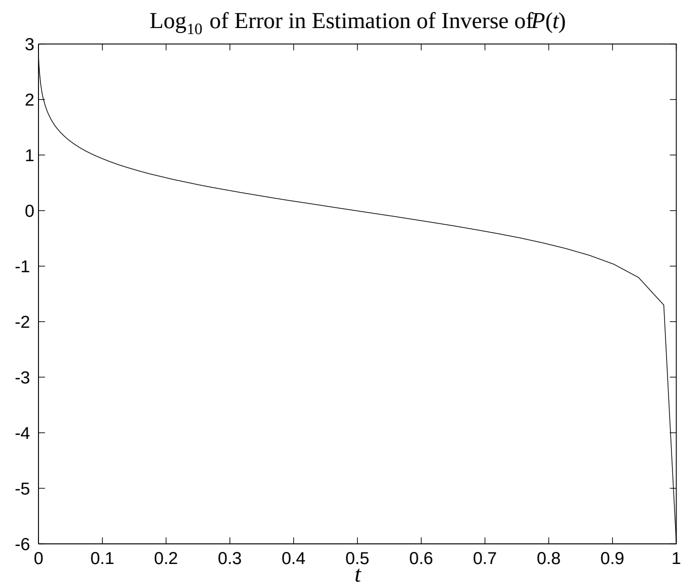


Figure 9: The base 10 log of the error $||\hat{x}(t)MM^T\hat{x}(t) - I||_\infty$, for Example 6.3.

In those same problems it may also be used for matrix inversion. For example, dynamic inversion has been incorporated into a controller for robotic manipulators in [25, 2] where the dynamic inverter produces inverse-kinematic solutions necessary for the control law. If inversion of, say, a time-varying mass matrix is also required in the same problem, a dynamic inverter may be augmented to provide that capability too, without interfering with other inversions within the same problem.

Acknowledgements

The authors are very grateful to C. A. Desoer for his thoughtful comments and suggestions.

References

[1] N. H. Getz and J. E. Marsden, "Dynamic inversion of nonlinear maps," Tech. Rep. 621, Center for Pure and Applied Mathematics, Berkeley, California, 19 December 1994.

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

[3] H. Rutishauser, "Ein infinitesimales analogon zum quotienten-differenzen-algorithmus," *Archiv der Mathematik*, 1954.

[4] H. Rutishauser, "Solution of eigenvalue problems with the lr-transformation," *National Bureau of Standards Applied Mathematics Series*, vol. 49, 1958.

[5] R. W. Brockett, "Dynamical systems that sort lists, diagonalize matrices, and solve linear programming problems," *Linear Algebra and Its Applications*, vol. 146, pp. 79–91, February 1991.

[6] M. Chu, "Scaled toda-like flows," *Linear Algebra and its Applications*, vol. 215, 15 January 1995.

[7] A. Bloch, "Estimation, principal components, and hamiltonian systems," *Systems and Control Letters*, vol. 6, 1985.

[8] A. Bloch, "Steepest descent, linear programming, and hamiltonian flows," *Contemporary Mathematics*, vol. 114, 1990.

[9] M. Chu, "Matrix differential equations - a continuous realization process for linear algebra problems," *Nonlinear Analysis-Theory Methods and Applications*, vol. 18, June 1992.

[10] M. Chu and K. Driessel, "Constructing symmetric nonnegative matrices with prescribed eigenvalues by differential equations," *SIAM Journal on Mathematical Analysis*, vol. 22, September 1991.

[11] S. T. Smith, "Dynamical systems that perform the singular value decomposition," *Systems and Control Letters*, vol. 16, no. 5, 1991.

[12] U. Helmke, J. B. Moore, and J. E. Perkins, "Dynamical systems that compute balanced realizations and the singular value decomposition," *Siam Journal of Matrix Analysis and Its Applications*, vol. 15, pp. 733–754, July 1994.

[13] 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.

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

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

[16] E. Allgower and K. Georg, "Simplical and continuation methods for approximating fixed points and solutions to systems of equations," *SIAM Review*, vol. 22, January 1980.

[17] J. P. Dunyak, J. L. Junkins, and L. T. Watson, "Robust nonlinear least squares estimations using the Chow-Yorke homotopy method," *Journal of Guidance*, vol. 7, pp. 752–755, 1984.

[18] L. T. Watson, "Engineering applications of the Chow-Yorke algorithm," *Applied Mathematical Computation*, vol. 9, pp. 111–133, 1981.

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

[20] J. Marsden and T. Hughes, *Mathematical foundations of elasticity.* Prentice-Hall Civil Engineering and Engineering Mechanics Series, Englewood Cliffs, N.J.: Prentice-Hall, 1983.

[21] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis.* New York: Cambridge University Press, 1991.

[22] R. Bartels and G. H. Stewart, "Algorithm 432, solution of the matrix equation AX+XB=C," *Communications of the ACM*, vol. 15, 1972.

[23] G. Golub, S. Nash, and C. V. Loan, "A Hessenberg-Scheer method for the problem AX+XB=C," *IEEE Transactions on Automatic Control*, vol. AC-24, 1979.

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

[25] N. H. Getz and J. E. Marsden, "Tracking implicit trajectories," in *IFAC Symposium on Nonlinear Control Systems Design*, (Tahoe City), International Federation of Automatic Control, 25-28 June 1995.