

Diagonalization of Time Varying Symmetric Matrices

Markus Baumann and Uwe Helmke

Mathematisches Institut, Universität Würzburg, Am Hubland, 97074 Würzburg
ms.b@gmx.de, helmke@mathematik.uni-wuerzburg.de

Abstract. We consider the task of diagonalizing symmetric time varying matrices $A(t)$. Based on the dynamic inversion technique developed by Getz and Marsden, a differential equation is proposed, whose solutions asymptotically track the diagonalizing transformation. In particular, one does not need to perfectly match the initial conditions, as the solutions converge exponentially towards the desired transformation. Thus, the proposed method is robust under perturbations.

1 Introduction

The aim of this paper is to establish an efficient algorithm for the diagonalization of time varying symmetric matrices. Thus we consider the task of determining orthogonal matrices $X(t)$ such that $X(t)'A(t)X(t)$ is diagonal. This is done by tracking a continuous time-dependent root of a suitable function $F_{A(t)}(X)$, whose roots $X_*(t)$ yield the desired solution. Therefore root finding methods, such as the dynamic inversion technique ([5]), can be applied to solve the problem. This approach leads to ordinary differential equations whose solutions converge exponentially fast to the diagonalizing transformations.

There exist already methods, which are able to track the diagonalizing transformation for a given symmetric matrix $A(t)$; see e.g. [1], [3], [4], [7].

These are given by the system of ODE's for D and X :

$$\dot{D} = X' \dot{A} X + D H - H D \quad (1)$$

$$\dot{X} = X H, \quad (2)$$

with H defined as: $H_{ij} = \begin{cases} 0, & i=j \\ \frac{(X' \dot{A} X)_{ij}}{a_j - a_i}, & i \neq j \end{cases}$

By solving this ODE, one obtains the diagonal matrix D and the corresponding orthogonal transformation matrix X . However, this algorithm only produces exact solutions, if it starts with perfect initial conditions. In particular, it does not perform any error-correction. In contrast, the dynamic inversion method achieves this purpose.

The paper is organized as follows. In the next section, a function $F(X, t)$ is introduced, which enables us to reformulate the diagonalization task as a root finding problem. Then, using the dynamic inversion technique, differential equations are proposed, that asymptotically compute the time-varying roots $X_*(t)$.

Examples are given in Section 3, which illustrate the convergence-properties of the new algorithms. Conclusions appear in Section 4.

2 Diagonalization of symmetric time-varying Matrices

Let $A(t) \in \mathbb{R}^{n \times n}$, $t \in \mathbb{R}$, be a continuously differentiable family of real symmetric matrices, with eigenvalues $\lambda_1(t), \dots, \lambda_n(t)$. $A(t)$ is supposed to satisfy the conditions:

1. $\|A(t)\|$ and $\|\dot{A}(t)\|$ are uniformly bounded on \mathbb{R} .
2. There exists a constant $m > 0$ such that $|\lambda_i - \lambda_j| \geq m$ ($i \neq j$) for all t .

Under these assumptions, there exists a continuously differentiable family of real orthogonal transformations $X_*(t)$, such that $X_*(t)'A(t)X_*(t)$ is diagonal for all $t \in \mathbb{R}$ ([6]). Our goal is to track such transformations by a suitable differential equation.

2.1 Reformulation of the problem

Consider the function

$$F : \mathbb{R}^{n \times n} \times \mathbb{R} \rightarrow \mathbb{R}^{n \times n} \quad (3)$$

defined by

$$F(X, t) = [N, X'A(t)X] + X'X - I, \quad (4)$$

where I is the identity matrix, $N = \text{diag}(1, \dots, n)$ and $[,]$ is the Lie-Bracket product defined as $[A, B] := AB - BA$ for $A, B \in \mathbb{R}^{n \times n}$.

Lemma 1. *$F(X, t) = 0$ if and only if X is an orthogonal matrix such that $X'A(t)X$ is diagonal.*

Proof. Note that the first summand of F is skew symmetric while the second one is symmetric. Thus F vanishes if and only if the two summands vanish, i.e. if and only if X is orthogonal and

$$[N, X'AX] = 0.$$

Since N is diagonal with distinct eigenvalues, the result follows.

Using this function F , the task of finding an orthogonal transformation X such that $X'A(t)X$ is diagonal, is equivalent to that of finding a root of $F(X, t)$. Thus, the eigenvalue and eigenvector-problem is reformulated as a root-finding-problem.

To track those solutions $X_*(t)$, we apply the technique of dynamic inversion. In order to do so, certain technical assumptions made in [5] have to be checked. This is done in the next lemma.

Lemma 2. *Let $A(t)$ be as above. There exists a continuously differentiable isolated solution $X_*(t)$ to $F(X, t) = 0$. $F(X, t)$ is C^∞ in X and C^1 in t . There exist constants $M, M_1 > 0$ such that*

- (i) $\|D_1^2 F(X_*(t), t)\| \leq M,$
- (ii) $\|D_1 F(X_*(t), t)^{-1}\| \leq M_1,$

holds for all $t \in \mathbb{R}$.

Proof. The claim concerning the differentiability properties of $F(X, t)$ is obvious.

The first and second partial derivatives of F w.r.to X are the linear and bilinear maps $D_1 F(X, t)$ and $D_1^2 F(X, t)$ respectively, given as

$$D_1 F(X, t) \cdot H = [N, H'AX + X'AH] + H'X + X'H, \tag{5}$$

$$D_1^2 F(X, t) \cdot (H, \hat{H}) = [N, H'A\hat{H} + \hat{H}'AH] + H'\hat{H} + \hat{H}'H, \tag{6}$$

and the partial derivative of $F(X, t)$ w.r.to t is

$$D_2 F(X, t) = [N, X'\dot{A}(t)X]. \tag{7}$$

From this we deduce the operator norm estimates

$$\|D_1 F(X, t)\| \leq 2(1 + 2a\|N\|)\|X\|, \tag{8}$$

and $\|D_1^2 F(X, t)\| \leq 2(1 + 2a\|N\|)$, where a denotes the infinity-norm of A .

In particular, $D_1^2 F(X, t)$ is uniformly bounded with respect to (X, t) . This shows (i).

We next show, that the partial derivative operator $D_1 F(X_*, t)$ is invertible for any solution (X_*, t) of $F(X, t) = 0$. In particular, X_* is orthogonal and

$$D_1 F(X_*, t) \cdot H = [N, H'AX_* + X_*'AH] + H'X_* + X_*'H. \tag{9}$$

Substituting $H = X_* \cdot \xi$, for $\xi \in \mathbb{R}^{n \times n}$ arbitrary, we obtain

$$\begin{aligned} D_1 F(X_*, t) \cdot (X_* \xi) &= [N, \xi' X_*' A X_* + X_*' A X_* \xi] + \xi' X_*' X_* + X_*' X_* \xi. \tag{10} \\ &= [N, \xi' D + D \xi] + \xi' + \xi, \end{aligned}$$

where $D = X_*' A X_*$ is diagonal. Thus $X_* \xi$ is in the kernel of $D_1 F(X_*, t)$ if and only if ξ is skew symmetric and $[N, [D, \xi]] = 0$. Hence $[D, \xi]$ must be diagonal and since D has distinct diagonal entries we conclude that $\xi = 0$. This shows that $D_1 F(X_*, t)$ is invertible for any root of F . By the implicit function theorem it follows that for every orthogonal X_0 with $X_0' A(0) X_0$ diagonal, there exists a unique C^1 -curve $X_*(t)$ of orthogonal matrices with $X_*(0) = X_0$. This shows the first claim.

To prove (ii), we derive a lower bound for the eigenvalues of $D_1 F(X_*, t)$. Let ξ_{pq} denote the entry of ξ with the largest absolute value. Assuming that the

norm of ξ is equal to one, the absolute value of ξ_{pq} is at least $\frac{1}{n^2}$. The smallest eigenvalue of $D_1F(X_*, t)$ is lower bounded by the sum of squares

$$(\xi_{pq}\lambda_p(p - q) + \xi_{qp}\lambda_q(p - q))^2 + (\xi_{pq} + \xi_{qp})^2$$

of the pq -entries of $[N, \xi'D + D\xi]$ and $\xi' + \xi$. For $p = q$ this is lower bounded by $\frac{4}{n^4}$, while otherwise it is lower bounded by $(\xi_{pq}\lambda_p + \xi_{qp}\lambda_q)^2 + (\xi_{pq} + \xi_{qp})^2$. The latter is a quadratic function in ξ_{qp} with minimum:

$$\frac{\xi_{pq}^2(\lambda_p - \lambda_q)^2}{1 + \lambda_q^2} \geq \frac{(\lambda_q - \lambda_p)^2}{n^4(1 + \lambda_q^2)}.$$

This is the desired lower bound for the eigenvalues of $D_1F(X_*, t)$. Thus (ii) follows with

$$M' = n^2 \max\left(\frac{1}{2}, \frac{1 + a}{m}\right).$$

The next result follows from a straight forward perturbation argument; see e.g. [6].

Proposition 1. *Let $X_*(t)$ be a continuously differentiable orthogonal transformation, such that $X'_*(t)A(t)X_*(t)$ is diagonal for all t . Let $X(t)$ converge to $X_*(t)$ exponentially as t goes to infinity. Then $X'(t)A(t)X(t) \rightarrow X'_*(t)A(t)X_*(t)$ exponentially for $t \rightarrow \infty$.*

Main Results

A direct consequence of the Dynamic Inversion Theorem with vanishing error ([5], Theorem 2.3.5), which is applicable due to Lemma 2, can now be formulated as follows.

Theorem 1. *Let $A(t)$ be as above and X_0 be any orthogonal transformation that diagonalizes $A(0)$. Let $X_*(t)$ be the unique continuously differentiable solution to $F(X, t) = 0$ with $X_*(0) = X_0$. For any $\mu > 0$ sufficiently large, and any $X(0)$ sufficiently close to X_0 , the solution $X(t)$ to*

$$D_1F(X, t)\dot{X} = -\mu([N, X'AX] + X'X - I) - [N, X'\dot{A}X]$$

converges exponentially to $X_(t)$ as t goes to infinity. I.e. there exist some $k, b > 0$ such, that $\|X(t) - X_*(t)\| \leq ke^{-bt}$. Moreover, any solution $X(t)$ is orthogonal for all t , provided $X(0)$ is orthogonal.*

Remark 1. The constant μ in the previous theorem influences the rate of convergence of $X(t)$ to $X_*(t)$.

One difficulty with the above approach is that the differential equation is in an implicit form. In the following we show how to circumvent this problem by designing suitable explicit forms.

2.2 Dynamic Inversion using matrix representations of $D_1F(X, t)^{-1}$

Recall the function

$$F(X, t) = [N, X'AX] + X'X - I. \tag{11}$$

and its partial derivative given by

$$D_1F(X, t) \cdot H = [N, H'AX + X'AH] + H'X + X'H. \tag{12}$$

Let $B(X, t) \in \mathbb{R}^{n^2 \times n^2}$ denotes the matrix representation of $D_1F(X, t)$. Thus to compute the inverse of $D_1F(X, t)$, we have to invert a $n^2 \times n^2$ matrix. This difficulty can be avoided by augmenting a differential equation for the inverse of B .

In the following theorem let $\Gamma(t)$ denote an approximation for the inverse of $B(X, t)$, i.e. of the matrix representation for $D_1F(X, t)^{-1}$.

Theorem 2. *Let*

$$E(\Gamma, X, t) = -\Gamma \text{vec}([N, X' \dot{A}(t)X]),$$

$$F^\Gamma(\Gamma, X, t) := B(X, t)\Gamma - I,$$

$$E^\Gamma(\Gamma, X, t) := -\Gamma \frac{d}{dt} B(X, t) \Big|_{\dot{X} = \text{vec}(E(\Gamma, X, t))} \Gamma.$$

Then the solution $(\Gamma(t), X(t))$ of the ODE

$$\begin{bmatrix} \dot{\Gamma} \\ \text{vec}(\dot{X}) \end{bmatrix} = -\mu \begin{bmatrix} \Gamma F^\Gamma(\Gamma, X, t) \\ \Gamma \text{vec}(F(X, t)) \end{bmatrix} + \begin{bmatrix} E^\Gamma(\Gamma, X, t) \\ E(\Gamma, X, t) \end{bmatrix}$$

converges exponentially to $(B(X_(t), t), X_*(t))$, assuming that $(\Gamma(0), X(0))$ is sufficient close to $(B(X_*(0), 0), X_*(0))$.*

2.3 Dynamic Inversion by solving the Sylvester equation

We derive an explicit formula for the Sylvester equation associated with D_1F .

$$D_1F(X, t) \cdot H = [N, H'AX + X'AH] + X'H + H'X =: K + Y, \tag{13}$$

where K, Y are given skew-symmetric and symmetric matrices, respectively.

Equation (13) is equivalent to the following equations

$$X'H + H'X = Y, \tag{14}$$

$$[N, H'AX + X'AH] = K. \tag{15}$$

According to [2], a general solution to (14) is given by

$$H = P'(Z + \frac{1}{2}Q'YQ)Q^{-1}. \tag{16}$$

Here Z is skew-symmetric and arbitrary, and P, Q are arbitrary matrices satisfying $PXQ = I$.

For X sufficiently close to the orthogonal matrix X_* , X^{-1} exists, and we can choose $P = X^{-1}$ and $Q = I$. Hence (16) is equivalent to

$$H = (X^{-1})'(Z + \frac{1}{2}Y),$$

where $Z = -Z'$ is the only remaining variable. We plug this equation for H into (15) and obtain

$$[N, Z'X^{-1}AX + X'A(X^{-1})'Z] = K + \frac{1}{2}[Y'X^{-1}AX + X'A(X^{-1})'Y, N] \tag{17}$$

We replace X^{-1} by the approximation X' and $X'AX$ by $D := \text{diag}(X'AX) =: \text{diag}(d_1, \dots, d_n)$. For $X = X_* + \epsilon$, these approximations are $O(\epsilon)$. With this, (17) is equivalent to

$$NZ'D + NDZ - Z'DN - DZN = K + [\frac{1}{2}(Y'D + DY), N],$$

which enables us to obtain Z_{ij} :

$$Z_{ij} = \frac{(K + [\frac{1}{2}(Y'D + DY), N])_{ij}}{n_i d_i - n_i d_j + d_j n_j - d_i n_j}.$$

Using simple algebraic manipulations we arrive at the following tracking algorithm. Note that the proposed algorithm differs from [3], [4] by an additional feedback term, which is necessary for the tracking property.

Theorem 3. *Let $Y = \mu(X'X - I)$, $D := \text{diag}(X'AX)$, $d_i := D_{ii}$ and $\Omega := \tilde{\Omega} + \hat{\Omega}$, where*

1. $\tilde{\Omega}_{ij} = \begin{cases} \frac{(X'AX)_{ij}}{d_i - d_j}, & i \neq j, \\ 0, & i = j. \end{cases}$
2. $\hat{\Omega}_{ij} = \begin{cases} \frac{(\mu X'AX - YD)_{ij}}{d_i - d_j}, & i \neq j, \\ \frac{1}{2}Y_{ij}, & i = j. \end{cases}$

The solution to

$$\dot{X} = -X\Omega$$

converges exponentially to $X_*(t)$, if $X(0)$ is sufficiently close to $X_*(0)$ and $\mu > 0$ is sufficiently large.

Solving this ODE is considerably cheaper than solving the coupled system of theorem 2, as there is no need to compute matrices of size $n^2 \times n^2$. Of course, this also has a stabilizing impact on the numerical aspects of the algorithm.

3 Examples

We consider some numerical examples to demonstrate the new approach. In all subsequent figures the solid lines represent the theoretically exact solutions, while the dashed lines show the solutions of the ODE of Theorem 2 computed via a standard Runge-Kutta method.

Example 1. Let $A(t) = \begin{pmatrix} 10\sin(t) & 2 \\ 2 & 3\cos(t) \end{pmatrix}$

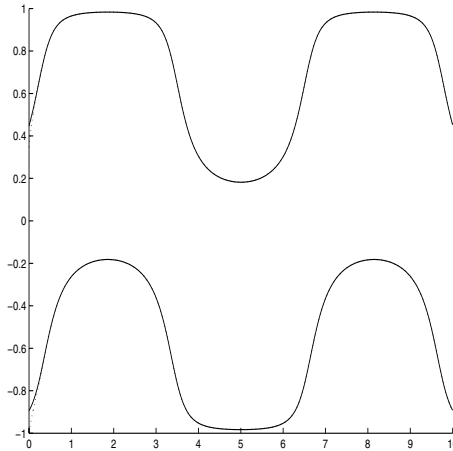


Fig. 1. Entries of the first column of $X(t)$ (dashed) and $X_*(t)$ (solid).

The figure shows a good approximation of the true eigenvectors/eigenvalues. This is to be expected, as the eigenvalues are well separated. In the next example, we will bring the eigenvalues very close together.

Example 2. Let $A(t) = \begin{pmatrix} t & 0.01 \\ 0.01 & 4 - t \end{pmatrix}$ This example shows, that the algorithm reacts sensitively, if the eigenvalues get too close. Nevertheless, after a short period of time, the algorithm is stabilizing again.

For the same example, but using the last algorithm from Theorem 3 instead of the first one, the sensitivity problems are eliminated. Thus the last algorithm appears to be preferable.

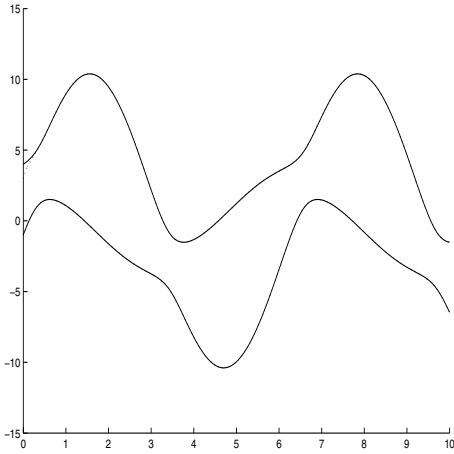


Fig. 2. Diagonal entries of $X(t)'A(t)X(t)$.

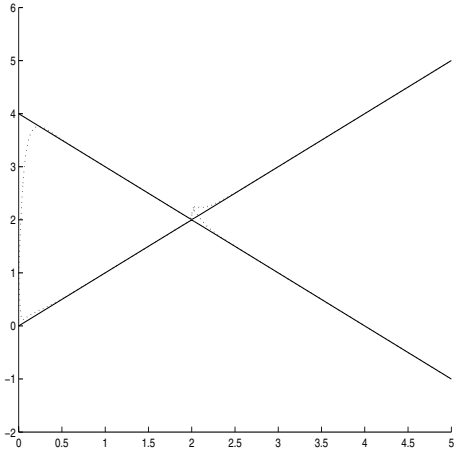


Fig. 3. Diagonal entries of $X(t)'A(t)X(t)$.

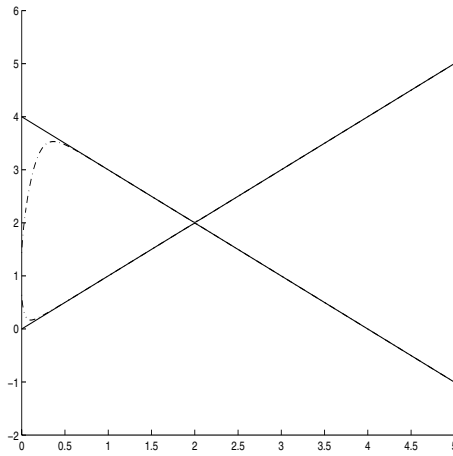


Fig. 4. Entries of the first column of $X(t)$ (dashed) and $X_*(t)$ (solid).

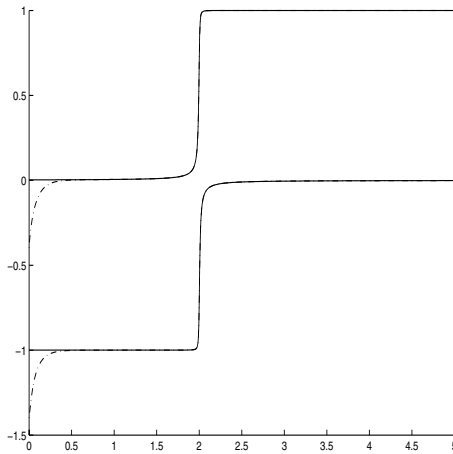


Fig. 5. Diagonal entries of $X(t)'A(t)X(t)$.

4 Conclusions

A new algorithm is proposed for tracking eigenvalues and eigenvectors of a time-varying symmetric matrix $A(t)$. The method depends on the technique of dynamic inversion previously introduced by Getz. A solution formula for Sylvester-equations leads to an explicit form of a differential equation that achieves robustly the tracking task.

5 Acknowledgement

Research partially supported by grants from the European Nonlinear Control Network (M.B.) and GIF I-526-34.6/1997 (U.H.).

References

- [1] Bell, S.J.G., Nichols, N.K.: A numerical method for the computation of the analytic singular value decomposition, Numerical Analysis Report 2, 1994.
- [2] Braden, H.W.: The equations $A'X \pm X'A = B$, SIAM Journal on Matrix Analysis and Applications, 20, (1998), 295–302.
- [3] Dehaene, J.: Continuous-Time Matrix Algorithms, Systolic Algorithms and Adaptive Neural Networks, Ph.D. Thesis KU Leuven, 1995.
- [4] Dieci, L., Eirola, T.: On smooth decompositions of matrices, SIAM Journal on Matrix Analysis and Applications, 20, (1999) 800-819.
- [5] Getz, N.H.: Dynamic Inversion of Nonlinear Maps, Ph.D. Thesis, Univ. of California at Berkeley, 1995.
- [6] Kato, T.: Perturbation Theory for Linear Operators, Springer-Verlag Berlin, 1966.
- [7] Wright, K.: Differential equations for the analytical singular value decomposition of a matrix, Numer. Math. 63:283 (1992).