JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 33, No. 1, JANUARY 1981

Variational Equations for the Eigenvalues and Eigenvectors of Nonsymmetric Matrices[1]

R. KALABA[2], K. SPINGARN[3], AND L. TESFATSION[4]

Abstract. The tracking of eigenvalues and eigenvectors for parameterized matrices is of major importance in optimization and stability problems. In the present paper, we consider a one-parameter family of matrices with distinct eigenvalues. A complete system of differential equations is developed for both the eigenvalues and the right and left eigenvectors. The computational feasibility of the differential system is demonstrated by means of a numerical example.

Key Words. Eigenvalues, eigenvectors, parameterized nonsymmetric matrices, variational equations.

1. Introduction: Basic Problem

Let $M(\alpha)$ be an $n \times n$ complex, matrix-valued differentiable function of a parameter α varying over some simply connected domain $\mathcal{C}^o$ of the complex plane $\mathcal{C}$. It will be assumed that $M(\alpha)$ has n distinct eigenvalues $\lambda_1(\alpha), \ldots, \lambda_n(\alpha)$ in $\mathcal{C}$ for each $\alpha \in \mathcal{C}^o$. Letting the superscript T denote transpose, it then follows (Ref. 1) that there exist two sets of linearly independent vectors

$$\{x_1(\alpha), \ldots, x_n(\alpha)\} \quad \text{and} \quad \{w_1(\alpha), \ldots, w_n(\alpha)\}$$

in $\mathcal{C}^n$ for each $\alpha \in \mathcal{C}^o$ satisfying

$$M(\alpha)x_i(\alpha) = \lambda_i(\alpha)x_i(\alpha), \qquad i = 1, \ldots, n, \tag{1a}$$

[1]The work of R. Kalaba and L. Tesfatsion was partially supported by the National Science Foundation under Grant No. ENG-77-28432 and by the National Institutes of Health under Grant No. GM-23732-03.

[2]Professor of Economics and Biomedical Engineering, University of Southern California, Los Angeles, California.
[3]Senior Staff Engineer, Space and Communications Group, Hughes Aircraft Company, Los Angeles, California.
[4]Assistant Professor of Economics, University of Southern California, Los Angeles, California.

0022-3239/81/0100-0001$03.00/0

$$M(\alpha)^T w_i(\alpha) = \lambda_i(\alpha) w_i(\alpha), \qquad i = 1, \ldots, n, \tag{1b}$$

$$x_i(\alpha)^T w_j(\alpha) \begin{cases} = 0, & \text{if } i \neq j, \\ \neq 0, & \text{if } i = j. \end{cases} \tag{1c}$$

The vectors

$$\{x_1(\alpha), \ldots, x_n(\alpha)\} \quad \text{and} \quad \{w_1(\alpha), \ldots, w_n(\alpha)\}$$

are generally referred to as the right and left eigenvectors of $M(\alpha)$, respectively.

The differentiability of $M(\alpha)$ over $\mathcal{C}^o$, coupled with the ruling out of exceptional points α in $\mathcal{C}^o$ where eigenvalues coalesce, guarantees (Ref. 2, Section II.1 and Sections II.4 to II.6) that the eigenvalues and right and left eigenvectors of $M(\alpha)$ have differentiable representations over $\mathcal{C}^o$. In this case, it is known (Ref. 2, p. 81) that

$$\dot{\lambda}_i(\alpha) = \frac{w_i(\alpha)^T \dot{M}(\alpha) x_i(\alpha)}{w_i^T(\alpha) x_i(\alpha)}, \qquad \alpha \in \mathcal{C}^o, \qquad i = 1, \ldots, n, \tag{2}$$

where a dot denotes differentiation with respect to α. However, corresponding analytical expressions for the derivatives $\dot{x}_i(\alpha)$ and $\dot{w}_i(\alpha)$ of the right and left eigenvectors of $M(\alpha)$ do not appear to be available in the literature. Without such additional equations, the system of differential equations (2) is analytically incomplete, in the sense that solutions $\lambda_i(\alpha)$ for (2) cannot be obtained by integration from initial conditions.

In the following section, the differential system (2) will be completed by providing differential equations for the right and left eigenvalues of $M(\alpha)$. As will be seen in Section 4 below, the resulting complete differential system provides a practical tool for numerical work.

2. Complete Variational Equations for Nonsymmetric Matrices

The exact form of the complete differential system for the eigenvalues and eigenvectors of $M(\alpha)$ depends on the normalization selected for the eigenvectors. We will start by imposing the general normalizations

$$x_i^T(\alpha) x_i(\alpha) = \varphi_i(\alpha), \qquad \alpha \in \mathcal{C}^o, \qquad i = 1, \ldots, n, \tag{3a}$$

$$w_i^T(\alpha) w_i(\alpha) = \Psi_i(\alpha), \qquad \alpha \in \mathcal{C}^o, \qquad i = 1, \ldots, n, \tag{3b}$$

for arbitrary differentiable functions

$$\varphi_i : \mathcal{C}^o \to \mathcal{C} \quad \text{and} \quad \Psi_i : \mathcal{C}^o \to \mathcal{C}.$$

Subsequently, (3) will be specialized to

$$\varphi_i(\alpha) \equiv \Psi_i(\alpha) \equiv 1,$$

and also to another convenient normalization.

Let $i \in \{1, \ldots, n\}$ and $\alpha \in \mathcal{C}^o$ be given. Differentiating (1a) with respect to α, and suppressing reference to α for ease of notation, one obtains

$$\dot{M}x_i + M\dot{x}_i = \dot{\lambda}_i x_i + \lambda_i \dot{x}_i. \tag{4}$$

Multiplying through (4) by w_i^T, we have

$$w_i^T \dot{M} x_i + w_i^T M \dot{x}_i = \dot{\lambda}_i w_i^T x_i + \lambda_i w_i^T \dot{x}_i. \tag{5}$$

Since

$$w_i^T M = \lambda_i w_i^T$$

by (1b), and since

$$w_i^T x_i \neq 0$$

by (1c), (5) reduces to

$$w_i^T \dot{M} x_i / w_i^T x_i = \dot{\lambda}_i, \tag{6}$$

the familiar differential equation for $\dot{\lambda}_i$.

Since the set of right eigenvalues $\{x_1, \ldots, x_n\}$ for $M(\alpha)$ spans $\mathcal{C}^n$, there exist coefficients β_{ij}, $j = 1, \ldots, n$, such that

$$\dot{x}_i = \sum_{j=1}^{n} \beta_{ij} x_j. \tag{7}$$

For $k \neq i$, it follows from 1(c) and (7) that

$$w_k^T \dot{x}_i = \sum_{j=1}^{n} \beta_{ij} w_k^T x_j = \beta_{ik} w_k^T x_k. \tag{8}$$

Hence, combining (7) and (8), we have

$$\dot{x}_i = \sum_{j \neq i} (w_j^T \dot{x}_i / w_j^T x_j) x_j + \beta_{ii} x_i. \tag{9}$$

Multiplying through (9) by x_i^T and solving for β_{ii}, we obtain

$$\beta_{ii} = (1/x_i^T x_i)[x_i^T \dot{x}_i - \sum_{j \neq i} (w_j^T \dot{x}_i / w_j^T x_j) x_i^T x_j]. \tag{10}$$

It is now necessary to replace the expressions $w_j^T \dot{x}_i$ and $x_i^T \dot{x}_i$ in (9) and (10) by a suitable expression independent of $\dot{x}_i$. Multiplying through (4) by w_j^T, $j \neq i$, we have

$$w_j^T \dot{M} x_i + w_j^T M \dot{x}_i = \dot{\lambda}_i w_j^T x_i + \lambda_i w_j^T \dot{x}_i. \tag{11}$$

By (1b) and (1c), (11) reduces to

$$w_j^T \dot{M} x_i + \lambda_j w_j^T \dot{x}_i = \lambda_i w_j^T \dot{x}_i. \tag{12}$$

Since the roots λ_j and λ_i are distinct by assumption, (12) can be restated as

$$w_j^T \dot{M} x_i / (\lambda_i - \lambda_j) = w_j^T \dot{x}_i. \tag{13}$$

Finally, the normalization (3a) implies that

$$\tfrac{1}{2}\dot{\varphi}_i = x_i^T \dot{x}_i. \tag{14}$$

Substituting (3a), (10), (13), (14) into (9) yields the desired differential equation for the right eigenvector x_i:

$$\dot{x}_i = \sum_{j\neq i} [w_j^T \dot{M} x_i/(\lambda_i - \lambda_j) w_j^T x_j] x_j + \left[\tfrac{1}{2}\dot{\varphi}_i - \sum_{j\neq i} [w_j^T \dot{M} x_i/(\lambda_i - \lambda_j) w_j^T x_j] x_i^T x_j\right] x_i/\varphi_i. \tag{15}$$

Equations analogous to (15) are similarly obtained for the left eigenvectors w_i. The complete differential system for the eigenvalues and right and left eigenvectors of $M(\alpha)$ thus has the form

$$\dot{\lambda}_i = w_i^T \dot{M} x_i / w_i^T x_i, \qquad i = 1, \ldots, n, \tag{16a}$$

$$\dot{x}_i = \sum_{j\neq i} [w_j^T \dot{M} x_i/(\lambda_i - \lambda_j) w_j^T x_j] x_j + [\tfrac{1}{2}\dot{\varphi}_i - \sum_{j\neq i} [w_j^T \dot{M} x_i/(\lambda_i - \lambda_j) w_j^T x_j] x_i^T x_j] x_i/\varphi_i, \; i = 1, \ldots, n, \tag{16b}$$

$$\dot{w}_i = \sum_{j\neq i} [x_j^T \dot{M} w_i/(\lambda_i - \lambda_j) x_j^T w_j] w_j + [\tfrac{1}{2}\dot{\Psi}_i - \sum_{j\neq i} [x_j^T \dot{M} w_i/(\lambda_i - \lambda_j) x_j^T w_j] w_i^T w_j] w_i/\Psi_i, \; i = 1, \ldots, n. \tag{16c}$$

Though our primary purpose is to use Eqs. (16) for computational purposes some analytical consequences are immediate. For example, for the Perron root of a positive matrix (see Ref. 5), Eq. (16a) implies that adding a positive matrix to a given positive matrix cannot decrease the Perron root.

System (16) is considerably simplified if the selected normalization is

$$\varphi_i(\alpha) \equiv \Psi_i(\alpha) \equiv 1, \qquad i = 1, \ldots, n.$$

In this case, (16) reduces to

$$\dot{\lambda}_i = w_i^T \dot{M} x_i / w_i^T x_i, \qquad i = 1, \ldots, n, \tag{17a}$$

$$\dot{x}_i = [I - x_i x_i^T] \sum_{j\neq i} [w_j^T \dot{M} x_i/(\lambda_i - \lambda_j) w_j^T x_j] x_j, \qquad i = 1, \ldots, n, \tag{17b}$$

$$\dot{w}_i = [I - w_i w_i^T] \sum_{j\neq i} [x_j^T \dot{M} w_i/(\lambda_i - \lambda_j) x_j^T w_j] w_j, \qquad i = 1, \ldots, n. \quad \text{(17c)}$$

An even greater simplification occurs if the $\varphi_i(\cdot)$ and $\Psi_i(\cdot)$ functions are selected so that the final terms in parentheses in (16b) and (16c) vanish identically. The system (16) then reduces to

$$\dot{\lambda}_i = w_i^T \dot{M} x_i / w_i^T x_i, \qquad i = 1, \ldots, n, \quad \text{(18a)}$$

$$\dot{x}_i = \sum_{j\neq i} [w_j^T \dot{M} x_i/(\lambda_i - \lambda_j) w_j^T x_j] x_j, \qquad i = 1, \ldots, n, \quad \text{(18b)}$$

$$\dot{w}_i = \sum_{j\neq i} [x_j^T \dot{M} w_i/(\lambda_i - \lambda_j) x_j^T w_j] w_j, \qquad i = 1, \ldots, n. \quad \text{(18c)}$$

At various stages in the integration of (18), the right and left eigenvectors can be normalized to unit length to prevent their magnitudes from becoming inconveniently large or small.

A numerical example in Section 4 illustrates the use of both (17) and (18).

3. Complete Variational Equations for Symmetric Matrices

Suppose that $M(\alpha)$ satisfies

$$M(\alpha) = M(\alpha)^T, \qquad \text{for each } \alpha \in \mathcal{C}^o.$$

In this case, the right and left eigenvectors of $M(\alpha)$ coincide over $\mathcal{C}^o$; hence, the normalizations (1c) and (3) together yield

$$x_i^T(\alpha) x_j(\alpha) = \begin{cases} 0, & \text{if } i \neq j, \\ \varphi_i(\alpha), & \text{if } i = j. \end{cases} \quad \text{(19)}$$

The differential system (16) thus reduces to

$$\dot{\lambda}_i = x_i^T \dot{M} x_i / \varphi_i, \qquad i = 1, \ldots, n, \quad \text{(20a)}$$

$$\dot{x}_i = \sum_{j\neq i} [x_j^T \dot{M} x_i/(\lambda_i - \lambda_j)\varphi_i] x_j + \tfrac{1}{2}(\dot{\varphi}_i/\varphi_i) x_i, \qquad i = 1, \ldots, n. \quad \text{(20b)}$$

If $\varphi_i(\alpha) \equiv 1$, $i = 1, \ldots, n$, (20) is further simplified to

$$\dot{\lambda}_i = x_i^T \dot{M} x_i, \qquad i = 1, \ldots, n, \quad \text{(21a)}$$

$$\dot{x}_i = \sum_{j\neq i} [x_j^T \dot{M} x_i/(\lambda_i - \lambda_j)] x_j, \qquad i = 1, \ldots, n. \quad \text{(21b)}$$

As noted in Ref. 2, p. 81, equations analogous to (21) for partial differential operators are familiar formulas in quantum mechanics. See, e.g., Ref. 3, Chapter 11, pp. 383–384. The differential form of (21a) is

$$d\lambda_i = x_i^T dMx_i. \tag{22}$$

It is known, by the Courant–Fischer minimax theorem (Refs. 4–5), that the addition of a positive definite matrix to a given positive definite matrix M increases all of the eigenvalues of M. This result is immediately obtainable from (22). For economic applications of Eqs. (21), see Ref. 6, pp. 240–246.

4. Illustrative Numerical Example

Consider a matrix-valued function $M(\alpha)$ defined over $\alpha \in R$ by

$$M(\alpha) \equiv \begin{bmatrix} 1 & \alpha \\ \alpha^2 & 3 \end{bmatrix}. \tag{23}$$

For this simple example, analytical expressions are easily obtainable for the eigenvalues $\{\lambda_1(\alpha), \lambda_2(\alpha)\}$ and the right and left eigenvectors $\{x_1(\alpha), x_2(\alpha), w_1(\alpha), w_2(\alpha)\}$ of $M(\alpha)$, where the latter are normalized to have unit length. Specifically,

$$\lambda_1(\alpha) = 2 + \gamma(\alpha), \tag{24a}$$

$$\lambda_2(\alpha) = 2 - \gamma(\alpha), \tag{24b}$$

$$x_1(\alpha) = [\alpha k_1/(1+\gamma(\alpha)), k_1]^T, \tag{24c}$$

$$x_2(\alpha) = [\alpha k_2/(1-\gamma(\alpha)), k_2]^T, \tag{24d}$$

$$w_1(\alpha) = [(\gamma(\alpha)-1)k_3/\alpha, k_3]^T, \tag{24e}$$

$$w_2(\alpha) = [(-\gamma(\alpha)-1)k_4/\alpha, k_4]^T, \tag{24f}$$

where

$$\gamma(\alpha) \equiv \sqrt{(1+\alpha^3)}$$

and the constants $k_1, \ldots, k_4$ are given by

$$k_1 \equiv 1/\sqrt{[\alpha^2/(1+\gamma(\alpha))^2+1]}, \tag{25a}$$

$$k_2 \equiv 1/\sqrt{[\alpha^2/(1-\gamma(\alpha))^2+1]}, \tag{25b}$$

$$k_3 \equiv 1/\sqrt{[(\gamma(\alpha)-1)^2/\alpha^2+1]}, \tag{25c}$$

$$k_4 \equiv 1/\sqrt{[(-\gamma(\alpha)-1)^2/\alpha^2+1]}. \tag{25d}$$

Note that the eigenvalues of $M(\alpha)$ are real iff $\alpha \geq -1.0$.

A numerical solution was first obtained for the eigenvalues and right and left eigenvectors of $M(\alpha)$ over the α interval [0.5, 2.0] by integrating

Table 1. Eigenvalues and eigenvectors for $\alpha = 2.0$.

	Numerical solution					
	Normalized Eqs. (18)		Unit normalized Eqs. (17)		Unit normalized analytical solution	
	Component		Component		Component	
	1	2	1	2	1	2
λ_1	5.0	—	5.0	—	5.0	—
λ_2	−1.0	—	−1.0	—	−1.0	—
x_1	0.449584	0.899168	0.447214	0.894427	0.447214	0.894427
x_2	−0.73633	0.73633	−0.707107	0.707107	−0.707107	0.707107
w_1	0.73633	0.73633	0.707107	0.707107	0.707107	0.707107
w_2	−0.899168	0.449584	−0.894427	0.447214	−0.894427	0.447214

the unit normalized differential equations (17). A fourth-order Runge–Kutta method was used for the integration with the α grid intervals set equal to 0.01. The integration was initialized by solving (24) for the eigenvalues and right and left eigenvectors of $M(\alpha)$ at $\alpha = 0.5$. As indicated in Table 1, the numerical solution obtained using the unit normalized differential equations (17) agreed with the analytical unit normalized solution (24) to at least six digits.

A numerical solution was also obtained for the eigenvalues and eigenvectors of $M(\alpha)$ over $[0.5, 2.0]$ by integrating the differential equations (18). By initializing the system as before, we were guaranteed that the magnitudes of the right and left eigenvectors $x_1, \ldots, x_n$ and $w_1, \ldots, w_n$ would be positive in some neighborhood of $\alpha = 0.5$, and in fact the magnitudes remained positive over the entire interval $[0.5, 2.0]$. As indicated in Table 1, the eigenvalues were obtained with the same accuracy as before. In addition, a subsequent unit normalization of the eigenvectors obtained via (18) yielded the same eigenvectors as obtained via the unit normalized differential equations (17).

As seen from (24a) and (24b), the eigenvalues of $M(\alpha)$ coalesce at $\alpha = -1.0$ and are complex for $\alpha < -1.0$. The unit normalized differential equations (17) are integrated from 0.5 to -1.0, using the integration stepsize of -0.01 for α. Six-digit accuracy for the eigenvalues and eigenvectors was obtained over $[-0.97, 0.5]$, degenerating to approximately two-digit accuracy at $\alpha = -1.0$. Similar results were obtained by integrating the differential equations (18) from 0.5 to -1.0.

5. Discussion

The present paper represents a first step toward the development of a computationally feasible procedure for tracking the eigenvalues and right and left eigenvectors of a parameterized matrix. Our main motivation has been the capability of modern-day computers to integrate, with great speed and accuracy, large-scale systems of ordinary differential equations subject to initial conditions. The computational feasibility of the initial value differential system developed in the present paper is illustrated by a numerical example.

In a subsequent paper, it will be shown that initial-value systems can also be developed for tracking a single eigenvalue together with one of its corresponding right or left eigenvectors.

References

1. GANTMACHER, F., *The Theory of Matrices, Vol. 1*, Chelsea Publishing Company, New York, New York, 1959.
2. KATO, T., *Perturbation Theory for Linear Operators*, Springer-Verlag, New York, New York, 1976.
3. KEMBLE, E., *The Fundamental Principles of Quantum Mechanics*, Dover Publications, New York, New York, 1958.
4. COURANT, R., and HILBERT, D., *Methoden der Mathematischen Physik*, Springer-Verlag, Berlin, Germany, 1931.
5. BELLMAN, R., *Introduction to Matrix Analysis*, McGraw-Hill Book Company, New York, New York, 1960.
6. THEIL, H., *Theory and Measurement of Consumer Demand*, Vol. 2, North-Holland Publishing Company, Amsterdam, Holland, 1976.