Econometric Theory, **1**, 1985, 179–191. Printed in the United States of Amercia.

ON DIFFERENTIATING EIGENVALUES AND EIGENVECTORS

JAN R. MAGNUS

London School of Economics

Let X_0 be a square matrix (complex or otherwise) and u_0 a (normalized) eigenvector associated with an eigenvalue λ_0 of X_0, so that the triple (X_0, u_0, λ_0) satisfies the equations $Xu = \lambda u$, $u_0^* u = 1$. We investigate the conditions under which unique differentiable functions $\lambda(X)$ and $u(X)$ exist in a neighborhood of X_0 satisfying $\lambda(X_0) = \lambda_0$, $u(X_0) = u_0$, $Xu = \lambda u$, and $u_0^* u = 1$. We obtain the first and second derivatives of $\lambda(X)$ and the first derivative of $u(X)$. Two alternative expressions for the first derivative of $\lambda(X)$ are also presented.

I. INTRODUCTION

The purpose of this paper is to provide explicit formulas (and simple proofs thereof) for the derivatives of eigenvalues and eigenvectors. These formulas are useful in the analysis of systems of dynamic equations and in many other applications. The somewhat obscure literature in this field (Lancaster [2], Neudecker [5], Sugiura [7], Bargmann and Nel [1], Phillips [6]) concentrates almost exclusively on the first derivative of an eigenvalue. Here we also obtain the derivative of the eigenvector and the second derivative of the eigenvalue.

The paper is organized as follows. In Section II we discuss two problems encountered in differentiating eigenvalues and eigenvectors, namely, the possible occurrence of complex or multiple eigenvalues. In Section III we obtain the first derivatives of eigenvalues and eigenvectors in the real symmetric case assuming simplicity of the eigenvalue. Section IV contains the complex analog. In Section V we obtain two alternative expressions for the first derivative of the eigenvalue function, and in Section VI we obtain the second derivative.

The following *notation* is used. Matrices are denoted by capital letters (A, X, Λ), vectors by lower-case roman letters (u, v), and scalars by lower-case Greek letters (λ, ε). An $m \times n$ matrix is one having m rows and n columns; A' denotes the transpose of A, A^+ its Moore–Penrose inverse, and

$r(A)$ its rank; if A is square, tr A denotes its trace, $|A|$ its determinant, and A^{-1} its inverse (when A is nonsingular); vec A denotes the column vector that stacks the columns of A one underneath the other, and $A \otimes B = (a_{ij}B)$ denotes the Kronecker product of A and B; $\mathbb{R}^{m \times n}$ is the class of real $m \times n$ matrices and $\mathbb{R}^n$ the class of real $n \times 1$ vectors, so $\mathbb{R}^n \equiv \mathbb{R}^{n \times 1}$. The $n \times n$ identity matrix is denoted I_n. The class of complex $m \times n$ matrices is denoted $\mathbb{C}^{m \times n}$. For $Z \in \mathbb{C}^{m \times n}$, $\bar{Z}$ denotes the complex conjugate of Z (i.e., the $m \times n$ matrix whose elements are the complex conjugates of the corresponding elements of Z), and Z^* denotes the $n \times m$ transpose of $\bar{Z}$.

II. TWO PROBLEMS

There are two problems involved in differentiating eigenvalues and eigenvectors. The first problem is that the eigenvalues of a real matrix A need not, in general, be real numbers—they may be complex. The second problem is the possible occurrence of multiple eigenvalues.

To appreciate the first point, consider the real 2×2 matrix function

$$A(\varepsilon) = \begin{bmatrix} 1 & \varepsilon \\ -\varepsilon & 1 \end{bmatrix}, \qquad \varepsilon \neq 0.$$

The matrix A is not symmetric, and its eigenvalues are $1 \pm i\varepsilon$. Since both eigenvalues are complex, the corresponding two eigenvectors must be complex as well; in fact, they are

$$\begin{bmatrix} 1 \\ i \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 \\ -i \end{bmatrix}.$$

We know, however, that if A is a *real symmetric* matrix, then its eigenvalues are real and its eigenvectors can always be taken to be real. Since the derivations in the real symmetric case are somewhat simpler, we begin our discussion by considering real symmetric matrices.

Thus, let X_0 be a real symmetric $n \times n$ matrix, and let u_0 be a (normalized) eigenvector associated with an eigenvalue λ_0 of X_0, so that the triple (X_0, u_0, λ_0) satisfies the equations

$$Xu = \lambda u, \qquad u'u = 1. \tag{1}$$

Since the $n + 1$ equations in (1) are implicit relations rather than explicit functions, we must first show that there exist explicit unique functions $\lambda(x)$ and $u(x)$ satisfying (1) in a neighborhood of X_0 and such that $\lambda(X_0) = \lambda_0$

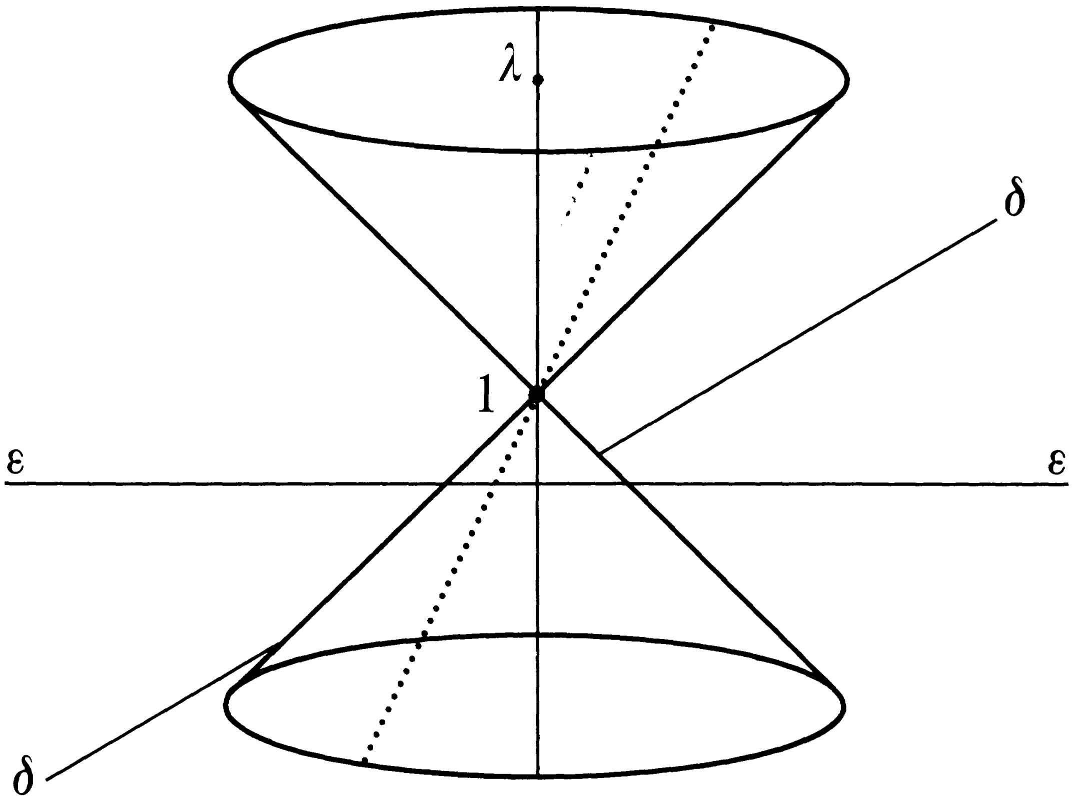

FIGURE 1.

and $u(X_0) = u_0$. Here the second (and more serious) problem arises—the possible occurrence of multiple eigenvalues.

We shall see that the implicit function theorem implies the existence of a neighborhood $N(X_0) \subset \mathbb{R}^{n \times n}$ of X_0, where the functions λ and u both exist and are ∞ times (continuously) differentiable, *provided* λ_0 *is a simple eigenvalue of* X_0. If, however, λ_0 is a multiple eigenvalue of X_0, then the conditions of the implicit function theorem are not satisfied. The difficulty is illustrated by the following example. Consider the real 2×2 matrix function

$$A(\varepsilon, \delta) = \begin{bmatrix} 1 + \varepsilon & \delta \\ \delta & 1 - \varepsilon \end{bmatrix}.$$

The matrix A is symmetric for every value of ε and δ; its eigenvalues are $\lambda_1 = 1 + (\varepsilon^2 + \delta^2)^{1/2}$ and $\lambda_2 = 1 - (\varepsilon^2 + \delta^2)^{1/2}$. Both eigenvalue functions are continuous in ε and δ, but clearly not differentiable at $(0, 0)$. (Strictly speaking we should also prove that λ_1 and λ_2 are the *only* two continuous eigenvalue functions.) The conical surface formed by the eigenvalues of $A(\varepsilon, \delta)$ has a singularity at $\varepsilon = \delta = 0$ (see Figure 1). For fixed ratio ε/δ, however, we can pass from one side of the surface to another going through $(0, 0)$ without noticing the singularity. This phenomenon is quite general and indicates the need of restricting the study of differentiability of multiple eigenvalues to one-dimensional perturbations only.[1]

III. THE REAL SYMMETRIC CASE

Let us now demonstrate the following theorem.

THEOREM 1. *Let* X_0 *be a real symmetric* $n \times n$ *matrix. Let* u_0 *be a normalized eigenvector associated with a simple eigenvalue* λ_0 *of* X_0. *Then a real-valued function* λ *and a vector function* u *are defined for all* X *in some neighborhood* $N(X_0) \subset \mathbb{R}^{n\times n}$ *of* X_0, *such that*

$$\lambda(X_0) = \lambda_0, \qquad u(X_0) = u_0, \tag{2}$$

and

$$Xu = \lambda u, \qquad u'u = 1, \qquad X \in N(X_0). \tag{3}$$

Moreover, the functions λ *and* u *are* ∞ *times differentiable on* $N(X_0)$, *and the differentials at* X_0 *are*

$$d\lambda = u_0'(dX)u_0 \tag{4}$$

and

$$du = (\lambda_0 I_n - X_0)^+(dX)u_0. \tag{5}$$

Equivalently, the derivatives at X_0 *are given by*

$$\frac{\partial\lambda}{\partial(\text{vec } X)'} = u_0' \otimes u_0' \tag{6}$$

(*or* $\partial\lambda/\partial X = u_0u_0'$), *and*

$$\frac{\partial u}{\partial(\text{vec } X)'} = u_0' \otimes (\lambda I_n - X_0)^+. \tag{7}$$

Note. In order for λ (and u) to be differentiable at X_0 we require that λ_0 is simple, but this does not, of course, exclude the possibility of multiplicities among the remaining $n-1$ eigenvalues of X_0.

Proof. Consider the vector function $f: \mathbb{R}^{n+1} \times \mathbb{R}^{n\times n} \to \mathbb{R}^{n+1}$ defined by the equation

$$f(u, \lambda; X) = \begin{bmatrix} (\lambda I_n - X)u \\ u'u - 1 \end{bmatrix},$$

and observe that f is ∞ times differentiable on $\mathbb{R}^{n+1} \times \mathbb{R}^{n \times n}$. The point $(u_0, \lambda_0; X_0)$ in $\mathbb{R}^{n+1} \times \mathbb{R}^{n \times n}$ satisfies

$$f(u_0, \lambda_0; X_0) = 0$$

and

$$\begin{vmatrix} \lambda_0 I_n - X_0 & u_0 \\ 2u_0' & 0 \end{vmatrix} \neq 0. \tag{8}$$

We note that the determinant in (8) is nonsingular if and only if the eigenvalue λ_0 is *simple* (in which case it takes the value of -2 times the product of the $n-1$ nonzero eigenvalues of $\lambda_0 I_n - X_0$).

The conditions of the implicit function theorem thus being satisfied, there exists a neighborhood $N(X_0) \subset \mathbb{R}^{n \times n}$ of X_0, a unique real-valued function $\lambda: N(X_0) \to \mathbb{R}$, and a unique (apart from its sign) vector function $u: N(X_0) \to \mathbb{R}^n$, such that

a. λ and u are ∞ times differentiable on $N(X_0)$,
b. $\lambda(X_0) = \lambda_0$, $u(X_0) = u_0$,
c. $Xu = \lambda u$, $u'u = 1$ for every $X \in N(X_0)$.

This completes the first part of our proof.

Let us now derive an explicit expression for $d\lambda$. From $Xu = \lambda u$ we obtain

$$(dX)u_0 + X_0\,du = (d\lambda)u_0 + \lambda_0\,du \tag{9}$$

where the differentials du and $d\lambda$ are defined at X_0. Premultiplying by u_0' gives

$$u_0'(dX)u_0 + u_0'X_0\,du = (d\lambda)u_0'u_0 + \lambda_0 u_0'\,du.$$

Since X_0 is symmetric we have $u_0'X_0 = \lambda_0 u_0'$. Hence

$$d\lambda = u_0'(dX)u_0,$$

because the eigenvector u_0 is normalized by $u_0'u_0 = 1$.

Next we derive an explicit expression for du. Let $Y_0 = \lambda_0 I - X_0$ and rewrite (9) as

$$Y_0\,du = (dX)u_0 - (d\lambda)u_0.$$

Premultiplying by Y_0^+ we obtain

$$Y_0^+ Y_0\,du = Y_0^+(dX)u_0,$$

because $Y_0^+ u_0 = 0$. (We use the fact that, for symmetric A, $Ab = 0$ iff $A^+ b = 0$.) To complete the proof we need only show that

$$Y_0^+ Y_0 \, du = du. \tag{10}$$

To prove (10) we note that $Y_0^+ u_0 = 0$ and $u_0' \, du = 0$ (because of the normalization $u'u = 1$). Hence

$$u_0'(Y_0^+ \; \vdots \; du) = 0'. \tag{11}$$

Since $u_0 \neq 0$, (11) implies that $r(Y_0^+ \; \vdots \; du) \leqslant n - 1$. Also, $r(Y_0^+) = r(Y_0) = n - 1$. It follows that

$$r(Y_0^+ \; \vdots \; du) = r(Y_0^+),$$

which is equivalent to (10).[2]

Finally, to obtain the derivatives, we rewrite (4) and (5) as

$$d\lambda = (u_0' \otimes u_0') \operatorname{vec} dX$$

and

$$du = [u_0' \otimes (\lambda_0 I - X_0)^+] \operatorname{vec} dX,$$

from which (6) and (7) follow. ■

Note 1. We have chosen to normalize the eigenvector u by $u'u = 1$, which means that u is a point on the unit ball. This is, however, not the only possibility. Another normalization,

$$u_0' u = 1, \tag{12}$$

though less common, is in many ways more appropriate. The reason for this will become clear when we discuss the complex case in the next section. If the eigenvectors are normalized according to (12), then u is a point in the hyperplane tangent (at u_0) to the unit ball. In either case we obtain $u' \, du = 0$ at $X = X_0$, which is all that is needed in the proof.

Note 2. It is important to note that, while X_0 is symmetric, the perturbations are not assumed to be symmetric. For symmetric perturbations, application of the chain rule immediately yields

$$d\lambda = (u_0' \otimes u_0') D \, dv(X), \qquad du = [u_0' \otimes (\lambda_0 I - X_0)^+] D \, dv(X), \tag{13}$$

and

$$\frac{\partial \lambda}{\partial [v(X)]'} = (u_0' \otimes u_0')D, \qquad \frac{\partial u}{\partial [v(X)]'} = [u_0' \otimes (\lambda_0 I - X_0)^+]D, \tag{14}$$

using the duplication matrix D and the $v(\cdot)$notation.[3]

IV. THE GENERAL COMPLEX CASE

Precisely the same techniques as were used in establishing Theorem 1 enable us to establish Theorem 2.

THEOREM 2. *Let* λ_0 *be a simple eigenvalue of a matrix* $Z_0 \in \mathbb{C}^{n \times n}$, *and let* u_0 *be an associated eigenvector, so that* $Z_0 u_0 = \lambda_0 u_0$. *Then a (complex) function* λ *and a (complex) vector function* u *are defined for all* Z *in some neighborhood* $N(Z_0) \in \mathbb{C}^{n \times n}$ *of* Z_0, *such that*

$$\lambda(Z_0) = \lambda_0, \qquad u(Z_0) = u_0, \tag{15}$$

and

$$Zu = \lambda u, \qquad u_0^* u = 1, \qquad Z \in N(Z_0). \tag{16}$$

Moreover, the functions λ *and* u *are* ∞ *times differentiable on* $N(Z_0)$, *and the differentials at* Z_0 *are*

$$d\lambda = v_0^*(dZ)u_0/v_0^* u_0 \tag{17}$$

and

$$du = (\lambda_0 I - Z_0)^+(I - u_0 v_0^*/v_0^* u_0)(dZ)u_0 \tag{18}$$

where v_0 *is the eigenvector associated with the eigenvalue* $\bar{\lambda}_0$ *of* Z_0^*, *so that* $Z_0^* v_0 = \bar{\lambda}_0 v_0$.

Note. It seems natural to normalize u by $v_0^* u = 1$ instead of $u_0^* u = 1$. Such a normalization does not, however, lead to a Moore–Penrose inverse in (18). Another possible normalization, $u^* u = 1$, also leads to trouble as the proof shows.

Proof. The facts that the functions λ and u exist and are ∞ times differentiable (i.e., analytic) in a neighborhood of Z_0 are proved in the same way as in Theorem 1. To find $d\lambda$ we differentiate both sides of $Zu = \lambda u$, and obtain

$$(dZ)u_0 + Z_0\, du = (d\lambda)u_0 + \lambda_0\, du, \tag{19}$$

where du and $d\lambda$ are defined at Z_0. We now premultiply by v_0^*, and since $v_0^* Z_0 = \lambda_0 v_0^*$ and $v_0^* u_0 \neq 0$, we obtain

$$d\lambda = v_0^*(dZ)u_0/v_0^* u_0.$$

To find du we define again $Y_0 = \lambda_0 I - Z_0$, and rewrite (19) as

$$\begin{aligned} Y_0\, du &= (dZ)u_0 - (d\lambda)u_0 \\ &= (dZ)u_0 - (v_0^*(dZ)u_0/v_0^* u_0)u_0 \\ &= (I - u_0 v_0^*/v_0^* u_0)(dZ)u_0. \end{aligned}$$

Premultiplying by Y_0^+ we obtain

$$Y_0^+ Y_0\, du = Y_0^+(I - u_0 v_0^*/v_0^* u_0)(dZ)u_0. \tag{20}$$

(Note that $Y_0^+ u_0 \neq 0$ in general.) To complete the proof we must again show that

$$Y_0^+ Y_0\, du = du. \tag{21}$$

From $Y_0 u_0 = 0$ we have $u_0^* Y_0^* = 0'$ and hence $u_0^* Y_0^+ = 0'$. Also, since u is normalized by $u_0^* u = 1$, we have $u_0^*\, du = 0$. (Note that $u^* u = 1$ does *not* imply $u_0^*\, du = 0$.) Hence

$$u_0^*(Y_0^+ \;\vdots\; du) = 0'.$$

As in the proof of Theorem 1, it follows that

$$r(Y_0^+ \;\vdots\; du) = r(Y_0^+),$$

which implies (21).[4] From (20) and (21), (18) follows. ■

V. TWO ALTERNATIVE EXPRESSIONS FOR $d\lambda$

As we have seen, the differential of the eigenvalue function associated with a simple eigenvalue λ_0 of a (complex) matrix Z_0 can be expressed as

$$d\lambda = \operatorname{tr} P_0\, dZ, \qquad P_0 = u_0 v_0^*/v_0^* u_0 \tag{22}$$

where u_0 and v_0 are (right and left) eigenvectors of Z_0 associated with λ_0:

$$Z_0 u_0 = \lambda_0 u_0, \qquad v_0^* Z_0 = \lambda_0 v_0^*, \qquad u_0^* u_0 = v_0^* v_0 = 1. \tag{23}$$

The matrix P_0 is idempotent with $r(P_0) = 1$.

Let us now express P_0 in two other ways, first as a product of $n-1$ matrices, then as a weighted sum of the matrices $I, Z_0, \dots, Z_0^{n-1}$.

THEOREM 3. *Let $\lambda_1, \lambda_2, \dots, \lambda_n$ be the eigenvalues of a matrix $Z_0 \in \mathbb{C}^{n\times n}$, and assume that λ_i is simple. Then a scalar function $\lambda_{(i)}$ exists, defined in a neighborhood $N(Z_0) \subset \mathbb{C}^{n\times n}$ of Z_0, such that $\lambda_{(i)}(Z_0) = \lambda_i$ and $\lambda_{(i)}(Z)$ is a (simple) eigenvalue of Z for every $Z \in N(Z_0)$. Moreover, $\lambda_{(i)}$ is ∞ times differentiable on $N(Z_0)$, and*

$$d\lambda_{(i)} = \operatorname{tr}\left[\left[\prod_{\substack{j=1\\ j\neq i}}^{n} \frac{\lambda_j I - Z_0}{\lambda_j - \lambda_j}\right] dZ\right]. \tag{24}$$

If we assume, in addition, that all *eigenvalues of Z_0 are simple, we may express $d\lambda_{(i)}$ also as*

$$d\lambda_{(i)} = \operatorname{tr}\left(\sum_{j=1}^{n} v^{ij} Z_0^{j-1}\, dZ\right), \qquad (i = 1, \dots, n), \tag{25}$$

where v^{ij} is the (i,j) element of the inverse of the Vandermonde matrix

$$V = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_n \\ \vdots & \vdots & & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \cdots & \lambda_n^{n-1} \end{bmatrix}.$$

Note. In expression (24) it is not demanded that the eigenvalues all be distinct, neither that they are all nonzero. In (25), however, the eigenvalues are assumed to be distinct. Still, one (but only one) eigenvalue may be zero.

Proof. Consider the following two matrices of order $n \times n$:

$$A = \lambda_i I - Z_0 \quad \text{and} \quad B = \prod_{j\neq i} (\lambda_j I - Z_0).$$

The Cayley–Hamilton theorem asserts that

$$AB = BA = 0. \tag{26}$$

Further, since λ_i is a simple eigenvalue of Z_0, the rank of A is $n - 1$. This, together with (26), implies

$$B = \mu u_0 v_0^* \tag{27}$$

where u_0 and v_0^* are defined in (23), and μ is an arbitrary scalar.

To determine the scalar μ, we use Schur's decomposition theorem and write

$$S^* Z_0 S = \Lambda + R, \qquad S^* S = I$$

Where Λ is a diagonal matrix containing $\lambda_1, \lambda_2, \ldots, \lambda_n$ on its diagonal, and R is *strictly* upper triangular. Then

$$\begin{aligned} \operatorname{tr} B &= \operatorname{tr} \prod_{j \neq i} (\lambda_j I - Z_0) = \operatorname{tr} \prod_{j \neq i} (\lambda_j I - \Lambda - R) \\ &= \operatorname{tr} \prod_{j \neq i} (\lambda_j I - \Lambda) = \prod_{j \neq i} (\lambda_j - \lambda_i). \end{aligned}$$

From (27) we also have

$$\operatorname{tr} B = \mu v_0^* u_0,$$

and since $v_0^* u_0$ is nonzero, we find

$$\mu = \frac{\prod_{j \neq i} (\lambda_j - \lambda_i)}{v_0^* u_0}.$$

Hence

$$\prod_{j \neq i} \frac{\lambda_j I - Z_0}{\lambda_j - \lambda_i} = \frac{u_0 v_0^*}{v_0^* u_0},$$

which, together with (22), proves (24).

Let us next prove (25). Since all eigenvalues of Z_0 are now assumed to be distinct, there exists a nonsingular matrix T such that

$$T^{-1} Z_0 T = \Lambda.$$

Therefore

$$\sum_j v^{ij} Z_0^{j-1} = T\left(\sum_j v^{ij} \Lambda^{j-1}\right) T^{-1}. \tag{28}$$

If we denote by E_{ii} the $n \times n$ matrix with 1 in its ith diagonal position and zeros elsewhere, and by δ_{ik} the Kronecker delta, then

$$\sum_j v^{ij}\Lambda^{j-1} = \sum_j v^{ij}\left(\sum_k \lambda_k^{j-1}E_{kk}\right) = \sum_k \left(\sum_j v^{ij}\lambda_k^{j-1}\right)E_{kk}$$
$$= \sum_k \delta_{ik}E_{kk} = E_{ii}, \tag{29}$$

because the scalar expression $\sum_j v^{ij}\lambda_k^{j-1}$ is the inner product of the ith row of V^{-1} and the kth column of V; that is,

$$\sum_j v^{ij}\lambda_k^{j-1} = \delta_{ik}.$$

Inserting (29) in (28) yields

$$\sum_j v^{ij}Z_0^{j-1} = TE_{ii}T^{-1} = (Te_i)(e_i'T^{-1}),$$

where e_i is the ith unit vector. Since λ_i is a simple eigenvalue of Z_0, we have

$$Te_i = \gamma u_0 \qquad \text{and} \qquad e_i'T^{-1} = \delta v_0^*$$

for some scalars γ and δ. Further,

$$1 = e_i'T^{-1}Te_i = \gamma\delta v_0^* u_0.$$

Hence

$$\sum_j v^{ij}Z_0^{j-1} = (Te_i)(e_i'T^{-1}) = \gamma\delta u_0 v_0^* = \frac{u_0 v_0^*}{v_0^* u_0}.$$

This, together with (22), proves (25). ■

VI. THE SECOND DIFFERENTIAL OF THE EIGENVALUE FUNCTION

As an application of the differential of the eigenvector du, let us now obtain $d^2\lambda$, the second differential of the eigenvalue function. We consider first the case in which X_0 is a real symmetric matrix.

THEOREM 4. *Under the same conditions as in Theorem 1, we have*

$$d^2\lambda = 2u_0'(dX)Y_0^+(dX)u_0 \tag{30}$$

where $Y_0 = \lambda_0 I_n - X_0$, *and the Hessian matrix is*

$$\frac{\partial^2 \lambda}{\partial \operatorname{vec} X \, \partial(\operatorname{vec} X)'} = K_n(Y_0^+ \otimes u_0 u_0' + u_0 u_0' \otimes Y_0^+) \tag{31}$$

where K_n *is the* $n^2 \times n^2$ *"commutation matrix."*[5]

Proof. Twice differentiating both sides of $Xu = \lambda u$, we obtain

$$2(dX)(du) + X_0 \, d^2u = (d^2\lambda)u_0 + 2(d\lambda)(du) + \lambda_0 \, d^2u$$

where all differentials are at X_0. Premultiplying by u_0' gives

$$d^2\lambda = 2u_0'(dX)(du), \tag{32}$$

since $u_0'u_0 = 1$, $u_0' \, du = 0$, and $u_0'X_0 = \lambda_0 u_0'$. From Theorem 1 we have $du = (\lambda_0 I - X_0)^+(dX)u_0$. Inserting this in (32) gives (30).

To prove (31), we rewrite (30) as

$$\begin{aligned} d^2\lambda &= 2(\operatorname{vec} dX')' \, (Y_0^+ \otimes u_0 u_0') \operatorname{vec} dX \\ &= 2(\operatorname{vec} dX)' K_n (Y_0^+ \otimes u_0 u_0') \operatorname{vec} dX, \end{aligned}$$

so that the (symmetric!) Hessian matrix is

$$\begin{aligned} \frac{\partial^2 \lambda}{\partial \operatorname{vec} X \, \partial(\operatorname{vec} X)'} &= K_n(Y_0^+ \otimes u_0 u_0') + (Y_0^+ \otimes u_0 u_0')K_n \\ &= K_n(Y_0^+ \otimes u_0 u_0' + u_0 u_0' \otimes Y_0^+). \end{aligned}$$

■

As in Note 2 of Section III we remark here that, while X_0 is symmetric, dX need not be symmetric. If we restrict X to be symmetric, we obtain for the Hessian matrix

$$\frac{\partial^2 \lambda}{\partial v(X) \partial [v(X)]'} = 2D'(u_0 u_0' \otimes Y_0^+)D, \tag{33}$$

using the fact that $KD = D$ (see Lemma 3.6 of [4]).

If λ_0 is the *largest* eigenvalue of X_0, then Y_0 is positive definite, and hence the Hessian matrix (33) is positive semidefinite. This corresponds to the fact that the largest eigenvalue is convex on the space of real symmetric matrices. Similarly, if λ_0 is the *smallest* eigenvalue of X_0, then the Hessian matrix is negative semidefinite, which corresponds to the fact that the smallest eigenvalue is concave on the space of real symmetric matrices.

The case in which Z_0 is a complex $n \times n$ matrix is proved in a similar way.

THEOREM 5. *Under the same conditions as in Theorem 2, we have*

$$d^2\lambda = \frac{2v_0^*(dZ)K_0(\lambda_0 I - Z_0)^+ K_0(dZ)u_0}{v_0^* u_0}, \tag{34}$$

where

$$K_0 = I - u_0 v_0^*/v_0^* u_0. \tag{35}$$

ACKNOWLEDGMENTS

I am very grateful to Roald Ramer for his many critical and helpful comments. The example and the figure in Section II were also his suggestions. I also thank Peter Phillips and Heinz Neudecker for their constructive remarks.

NOTES

1. Lancaster [2] has an interesting collection of results about the multiple-eigenvalue case with one-dimensional perturbations.
2. A more geometric proof of (10) would use the fact that the column space of Y_0^+ is orthogonal to the null space of Y_0.
3. For a definition and properties of the duplication matrix and the $v(\cdot)$ notation, see [4].
4. The remark of note 2 applies here as well.
5. For a definition and properties of the commutation matrix, see [3].

REFERENCES

1. Bargmann, R. E. and D. G. Nel. On the matrix differentiation of the characteristic roots of matrices. *South African Statistical Journal* 8 (1974): 135–144.
2. Lancaster, P. On eigenvalues of matrices dependent on a parameter. *Numerische Mathematik* 6 (1964): 377–387.
3. Magnus, J. R. and H. Neudecker. The commutation matrix: Some properties and applications. *The Annals of Statistics* 7 (1979): 381–394.
4. Magnus, J. R. and H. Neudecker. The elimination matrix: Some lemmas and applications. *SIAM Journal on Algebraic and Discrete Methods* 1 (1980): 422–449.
5. Neudecker, H. On matrix procedures for optimizing differentiable scalar functions of matrices. *Statistica Neerlandica* 21 (1967): 101–107.
6. Phillips, P.C.B. A simple proof of the latent root sensitivity formula. *Economics Letters* 9 (1982): 57–59.
7. Sugiura, N. Derivatives of the characteristic root of a symmetric or a hermitian matrix with two applications in multivariate analysis. *Communications in Statistics* 1 (1973): 393–417.