Journal of Research of the National Bureau of Standards Vol. 51, No. 1, July 1953 Research Paper 2429

Eigenvectors of Matric Polynomials [1]

Murray Mannos

It is the object of this paper to compare the eigenvectors of an arbitrary $n\times n$ matrix A over the complex field with those of the matric polynomial $f(A)$. While it is well known that each eigenvector of A is an eigenvector of $f(A)$, it is not, in general, true that A and $f(A)$ have identical eigenvectors. In this regard a necessary and sufficient condition that A and $f(A)$ have identical eigenvectors is given. The condition is that both (1) and (2) hold:

(1) $f'(\lambda)\neq 0$ for all eigenvalues λ of the matrix A corresponding to nonlinear elementary divisors.

(2) The values of $f(\mu)$ are distinct for all eigenvalues μ of the matrix A corresponding to linear elementary divisors.

When either (1) or (2) fails to hold, then $f(A)$ has eigenvectors that are not eigenvectors of A. This situation is also discussed.

The vector space of eigenvectors of the matrix A corresponding to the eigenvalue λ shall be denoted by $V_\lambda[A]$. The vector space spanned by the eigenvectors of A shall be denoted by $V[A]$. Further, $d\{V_\lambda[A]\}$ and $d\{V[A]\}$ shall denote their dimensions. It is clear that each eigenvector of A is an eigenvector of $f(A)$. That is, $V_\lambda[A]\subseteq V_{f(\lambda)}[f(A)]$ for each eigenvalue λ of A. Thus $V[A]\subseteq V[f(A)]$.

Let J be the Jordan canonical form of A. Then there exists a nonsingular matrix P such that $P^{-1}AP=J$ and so $P^{-1}f(A)P=f(J)$.

Lemma 1. *The eigenvectors of A and $f(A)$ are identical if, and only if, the eigenvectors of J and $f(J)$ are identical.*

Proof. This follows from the fact that $P\xi$ is an eigenvector of PBP^{-1} if ξ is an eigenvector of the matrix B.

Since $P\{V_\lambda[J]\}=V_\lambda[A]$, it follows that

$$d\{V[f(A)]\}-d\{V[A]\}=d\{V[f(J)]\}-d\{V[J]\},$$

where $P\{V_\lambda[J]\}$ denotes the space of all vectors $P\xi$, where ξ is an eigenvector of J corresponding to λ.

Lemma 2. *If $D=diag\ [\alpha,\alpha,\ldots,\alpha;\ \beta,\beta,\ldots,\beta;\ \ldots;\ \pi,\pi,\ldots,\pi]$, where $\alpha,\beta,\ldots,\pi$ are distinct, then $V[D]=V[f(D)]$. Furthermore, the eigenvectors of D are identical with those of $f(D)$ if, and only if, $f(\alpha), f(\beta),\ldots, f(\pi)$ are all distinct.*

Proof. If D is of order n, it is clear that both the eigenvectors of D and $f(D)$ each generate the whole n-dimensional vector space. If $f(\alpha)=f(\beta)$, then it is easily seen that $f(D)$ has eigenvectors corresponding to the eigenvalue $f(\alpha)=f(\beta)$, which are not eigenvectors of D.

Lemma 3. *If*

$$J_m(\lambda)=\begin{bmatrix} \lambda & 1 & 0 & \ldots & 0 & 0 & 0\\ 0 & \lambda & 1 & \ldots & 0 & 0 & 0\\ \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot\\ 0 & 0 & 0 & \ldots & 0 & \lambda & 1\\ 0 & 0 & 0 & \ldots & 0 & 0 & \lambda \end{bmatrix},$$

then

$$f(J_m(\lambda))=\begin{bmatrix} f(\lambda) & f'(\lambda) & \frac{f''(\lambda)}{2!} & \frac{f'''(\lambda)}{3!} & \ldots & & \frac{f^{(m-1)}(\lambda)}{(m-1)!}\\ 0 & f(\lambda) & f'(\lambda) & \frac{f''(\lambda)}{2!} & \ldots & & \frac{f^{(m-2)}(\lambda)}{(m-2)!}\\ 0 & 0 & f(\lambda) & f'(\lambda) & \ldots & & \frac{f^{(m-3)}(\lambda)}{(m-3)!}\\ \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot\\ \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot\\ \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot\\ 0 & 0 & 0 & 0 & \ldots & f(\lambda) & f'(\lambda)\\ 0 & 0 & 0 & 0 & \ldots & 0 & f(\lambda) \end{bmatrix}.$$

Proofs of the lemma may be found in Wedderburn [2] and MacDuffee.[3]

Lemma 4. *If $f'(\lambda)=f''(\lambda)=\ldots=f^{(r)}(\lambda)=0$ and $f^{(r+1)}(\lambda)\neq 0$ $(r=1,2,\ldots,m-2)$, then*

$$d\{V_{f(\lambda)}[f(J_m(\lambda))]\}-d\{V_\lambda[J_m(\lambda)]\}=r,$$

and conversely. Also,

$$d\{V_{f(\lambda)}[f(J_m(\lambda))]\}-d\{V_\lambda[J_m(\lambda)]\}=m-1$$

if, and only if, $f'(\lambda)=f''(\lambda)=\ldots=f^{m-1}(\lambda)=0$.

Proof. If $f'(\lambda)=f''(\lambda)=\ldots=f^{(r)}(\lambda)=0$, but $f^{(r+1)}(\lambda)\neq 0$, it follows that the rank of the matrix $f(J_m(\lambda))-f(\lambda)I$ is $m-(r+1)$, as may be observed by noting the nonzero minor of order $m-(r+1)$ in the upper right-hand corner of the matrix. Thus the nullity of $f(J_m(\lambda))-f(\lambda)I$ is $r+1$. If the first m-1 derivatives of $f(x)$ vanish at $x=\lambda$, then the nullity of $f(J_m(\lambda))-f(\lambda)I$ is m. In a similar fashion one finds that the matrix $J_m(\lambda)-\lambda I$ has nullity equal to 1. This establishes the lemma.

Corollary. The matrices $J_m(\lambda)$ and $f(J_m(\lambda))$ have identical eigenvectors if, and only if, $f'(\lambda)\neq 0$.

It is to be noted that in this case $V_{f(\lambda)}[f(J_m(\lambda))]$

[1] Preparation of this paper was sponsored (in part) by the Office of Scientific Research, USAF.

[2] J. H. M. Wedderburn, Lectures on Matrices, (Am. Math. Soc., Colloquium Pub. vol. 17, New York, N. Y., 1934).

[3] C. C. MacDuffee, An Introduction to Abstract Algebra (John Wiley & Sons, New York, N. Y., 1940).

and $V_\lambda[J_m(\lambda)]$ are each spanned by the single column vector $[1,0, \ldots, 0]^T$.

At this point it will be convenient to introduce the language of elementary divisors. Each block in the Jordan canonical form of a matrix A corresponds to an elementary divisor $(x-\lambda)^m$, and conversely. Each such block is called the hypercompanion matrix $J_m(\lambda)$ of the polynomial $(x-\lambda)^m$.

Suppose the nonlinear elementary divisors of the matrix A corresponding to the eigenvalue λ are of the form $(x-\lambda)^{m_\lambda^{(1)}}, (x-\lambda)^{m_\lambda^{(2)}}, \ldots, (x-\lambda)^{m_\lambda^{(k_\lambda)}}$, where $m_\lambda^{(i)}$ $(i=1,2, \ldots, k_\lambda)$ are integers such that $m_\lambda^{(1)}>m_\lambda^{(2)}>\ldots>m_\lambda^{(k_\lambda)}>1$. Furthermore, suppose that $(x-\lambda)^{m_\lambda^{(i)}}$ appears as an elementary divisor $n_\lambda^{(i)}$ times. Let $p_\lambda=\sum_{i=1}^{k_\lambda} n_\lambda^{(i)}$ denote the total number of nonlinear elementary divisors of A corresponding to the eigenvalue λ. Denote by n_λ the number of linear elementary divisors of A corresponding to the eigenvalue λ. Further, denote by K_λ the direct sum of the hypercompanion matrices of the elementary divisors corresponding to λ. Set

$$d_\lambda=d\{V_{f(\lambda)}[f(K_\lambda)]\}-d\{V_\lambda[K_\lambda]\}.$$

Lemma 5. *If the first $m_\lambda^{(1)}-1$ derivatives of $f(x)$ vanish at $x=\lambda$, then $d_\lambda=\sum_{i=1}^{k_\lambda} n_\lambda^{(i)}(m_\lambda^{(i)}-1)$; whereas, if the first $r<m_\lambda^{(1)}-1$ derivatives of $f(x)$ vanish at $x=\lambda$, but $f^{(r+1)}(\lambda)\neq 0$, then $d_\lambda=\sum_{i=1}^{l-1} n_\lambda^{(i)}r+\sum_{i=l}^{k_\lambda} n_\lambda^{(i)}(m_\lambda^{(i)}-1)$, where $m_\lambda^{(l)}$ is the largest of the integers $m_\lambda^{(i)}$ $(i=1,2,\ldots,k_\lambda)$ such that $m_\lambda^{(l)}\leq r$.*

Proof. This follows from lemma 4, the fact that the nullity of $K_\lambda-\lambda I$ is the sum of the nullities of the characteristic matrices of the hypercompanion matrices of the individual elementary divisors corresponding to the eigenvalue λ and a similar statement about the nullity of $f(K_\lambda)-f(\lambda)I$.

Corollary. The matrices K_λ and $f(K_\lambda)$ have identical eigenvectors if, and only if, $f'(\lambda)\neq 0$.

Proof. This follows from the fact that $d_\lambda=0$.

The Jordan canonical form $J=\text{diag}\,[K_\alpha,K_\beta,\ldots,K_\pi]$ of A is a direct sum of matrices K_λ, where $\lambda=\alpha,\beta,\ldots$ runs through the distinct eigenvalues of A. The subsequent theorems of this paper involve the following main conditions:

Condition 1. *$f'(\lambda)\neq 0$ for all eigenvalues λ of the matrix A corresponding to nonlinear elementary divisors.*

Condition 2. *The values $f(\mu)$ are distinct for all eigenvalues μ of the matrix A corresponding to linear elementary divisors.*

The first theorem concerns the case where condition 1 does not hold, whether condition 2 holds or not.

Theorem 1. *If condition* 1 *does not hold, then*

$$d\{V[f(A)]\}-d\{V[A]\}=\sum_\lambda d_\lambda,$$

where λ varies through all distinct eigenvalues of A corresponding to nonlinear elementary divisors, and d_λ is computed as in lemma 5.

Proof. One notes first that

$$V[J]=V_\alpha[J]+V_\beta[J]+\ldots+V_\pi[J],$$

and

$$V[f(J)]=V_{f(\alpha)}[f(J)]+V_{f(\beta)}[f(J)]+\ldots+V_{f(\pi)}[f(J)].$$

If $f(\alpha), f(\beta), \ldots, f(\pi)$ are distinct, for a fixed eigenvalue λ, the nullity of $J-\lambda I$ is the same as the nullity of $K_\lambda-\lambda I$, and the nullity of $f(J)-f(\lambda)I$ is the same as the nullity of $f(K_\lambda)-f(\lambda)I$. It follows that

$$d\{V_{f(\lambda)}[f(J)]\}-d\{V_\lambda[J]\}=d_\lambda.$$

Summing over distinct eigenvalues, one obtains

$$d\{V[f(J)]\}-d\{V[J]\}=\sum_\lambda d_\lambda.$$

By the statement following the proof of lemma 1 it follows that

$$d\{V[f(A)]\}-d\{V[A]\}=\sum_\lambda d_\lambda.$$

If $f(\alpha)=f(\beta)=\ldots=f(\rho)$, then the nullity of $f(J)-f(\alpha)I$ minus the sum of the nullities of $J-\alpha I$, $J-\beta I, \ldots$, and $J-\rho I$ is $d_\alpha+d_\beta+\ldots+d_\rho$ and

$$d\{V_{f(\alpha)}[f(J)]\}-d\{V_\alpha[J]\}-d\{V_\beta[J]\}-\ldots-d\{V_\rho[J]\}=d_\alpha+d_\beta+\ldots+d_\rho.$$

Summing, one obtains in either case

$$d\{V[f(A)]\}-d\{V[A]\}=\sum_\lambda d_\lambda,$$

where λ runs through all the distinct eigenvalues of A corresponding to nonlinear elementary divisors.

The next theorem concerns the case in which condition 1 holds but condition 2 does not.

Theorem 2. *If condition* 1 *holds, than $V[A]=V[f(A)]$.*

In this case $\sum_\lambda d_\lambda=0$, where λ varies over all distinct eigenvalues corresponding to nonlinear elementary divisors.

If $f(\alpha)=f(\beta)=\ldots=f(\rho)$, suppose that

$$d\{V_{f(\alpha)}[f(J)]\}=d\{V_\alpha[J]\}+d\{V_\beta[J]\}+\ldots+d\{V_\rho[J]\}=s.$$

Then as in lemma 2 the s-dimensional space $V_\alpha[J]+V_\beta[J]+\ldots+V_\rho[J]$ contains vectors that are not eigenvectors of J, but $V_{f(\alpha)}[f(J)]$ consists solely of eigenvectors of $f(J)$ corresponding to $f(\alpha)$.

Corollary. If $f(\alpha)=f(\beta)=\ldots=f(\rho)$, then

$$d\{V_{f(\alpha)}[f(A)]\}-d\{V_\alpha[A]\}=\sum_\lambda d_\lambda+(p_\lambda+n_\lambda),$$

where λ runs over all eigenvalues among $\alpha, \beta, \ldots, \rho$ which correspond to nonlinear elementary divisors and *where $p_\lambda + n_\lambda$ is the total number of elementary divisors (nonlinear and linear) corresponding to λ.*

Finally, the next result covers the case in which conditions 1 and 2 both hold.

Theorem 3. The matrices A and $f(A)$ have identical eigenvectors if, and only if (1) $f'(\lambda)\neq 0$ for all eigenvalues λ of the matrix A corresponding to nonlinear elementary divisors, and (2) the values $f(\mu)$ are distinct for all eigenvalues μ of the matrix A corresponding to linear elementary divisors.

Remark 1. It may readily be seen that $V[J]$ and $V[f(J)]$ each can be generated by a set of linearly independent vectors, each of which has a 1 in a single component and 0 in all the remaining components. Thus if $J=P^{-1}AP$ is the Jordan canonical form of A, than $V[f(A)]$ can be generated by a subset of the column vectors of P.

Remark 2. A simple application of the foregoing theory shows that if A is a 2×2 matrix, then A and $f(A)$ have the same eigenvectors unless either (1) A is diagonable and has distinct eigenvalues α, β for which $f(\alpha)=f(\beta)$, or (2) A is nondiagonable and $f(x)=k(x)(x-\alpha)^2+c$, where α is the eigenvalue of A, $k(x)$ is an arbitrary polynomial, and c is an arbitrary constant.

The following example is given as an illustration:

$$\text{Let } A=\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & -1 & -1 & -1 & -1 & -2 & -1 \\ 1 & 0 & 1 & 1 & 0 & -1 & 0 \end{bmatrix}$$

and $f(x)=x^6+x^5-x^4-x^3+2$.

$$\text{For } P=\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

it may be verified that

$$P^{-1}=\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ -1 & -1 & -1 & -1 & -1 & -1 & 1 \end{bmatrix}$$

and that the Jordan canonical form of A is

$$J=P^{-1}AP=\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix},$$

where

$$\begin{cases} K_1=J_1(1)=[1], \\ K_0=\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad J_3(0)=\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad J_1(0)=[0], \\ K_{-1}=J_2(-1)=\begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix}. \end{cases}$$

Since

$$\begin{cases} f(0)=f(1)=f(-1)=2, \\ f'(0)=f'(-1)=0, \quad f'(1)\neq 0, \\ f''(0)=0, \end{cases}$$

it follows that

$$f(J)=J^6+J^5-J^4-J^3+2I=\begin{bmatrix} 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix}$$

It is readily seen that $V_1[J]$ is generated by $[1,0,0,0,0,0,0]^T$, $V_0[J]$ is generated by $[0,1,0,0,0,0,0]^T$ and $[0,0,0,0,1,0,0]^T$, and $V_{-1}[J]$ is generated by $[0,0,0,0,0,1,0]^T$. Since $V[J]=V_1[J]+V_0[J]+V_{-1}[J]$, it follows that $d\{V[J]\}=4$. Since $f[J]$ is a scalar matrix, $d\{V[f(J)]\}=7$. Hence $d\{[f(A)]\}-d\{V[A]\}=3$. The same result is arrived at by the use of theorem 1, where $\sum_\lambda d_\lambda$ is calculated as in lemma 5. Since $f'(0)=0$, $f''(0)=0$, and $f'(-1)=0$, where 0, -1 are the eigenvalues corresponding to nonlinear elementary divisors, it follows that

$$d_0=n_0^{(1)}(m_0^{(1)}-1)=1\cdot 2=2$$

and

$$d_{-1}=n_1^{(1)}r=1\cdot 1=1$$

and

$$\sum_{\lambda} d_\lambda = d_0 + d_{-1} = 2+1=3,$$

as before. By the corollary following theorem 2,

$$d\{V_2[f(A)]\} - d\{V_1[A]\} = d_0 + d_{-1} + p_0 + p_{-1} + n_0 + n_{-1}$$
$$= 2+1+1+1+1+0=6,$$

which checks with the observed results above.

From a glance at the above eigenvectors generating $V[J]$, it is clear that $V[A]$ is generated by the first, second, fifth, and sixth-column vectors of P. Since $f(A)$ is a scalar, $V[f(A)]$ is generated by all seven column vectors of P.

The author thanks A. H. Clifford, O. Taussky, and N. Wiegmann for their helpful suggestions.

WASHINGTON, December 1, 1952.