Zeitschrift: L'Enseignement Mathématique

Herausgeber: Commission Internationale de l'Enseignement Mathématique

Band: 52 (2006)

Heft: 3-4: L'enseignement mathématique

Artikel: Interpolation approach to the spectral resolution of square matrices

Autor: Verde-Star, Luis

DOI: https://doi.org/10.5169/seals-2234

L'Enseignement Mathématique (2) 52 (2006), 239–253

INTERPOLATION APPROACH TO THE SPECTRAL RESOLUTION OF SQUARE MATRICES

by Luis VERDE-STAR

ABSTRACT. We present a proof of the spectral resolution theorem for square matrices that are not necessarily diagonalizable. The construction of the idempotent and nilpotent component matrices and the proofs of their properties use only simple properties of the basic Hermite interpolation polynomials. The relevant results from polynomial interpolation are presented in detail. Determinants, canonical forms, inner products, and integrals are not used in our development.

1. INTRODUCTION

The spectral decomposition theorem for linear operators on finite dimensional spaces is a very important result. Its generalizations to infinite dimensions constitute a fundamental part of the theory of operators. The spectral resolution may be used in many situations as an alternative to the Jordan canonical form, since it gives a decomposition of a linear operator as a sum of orthogonal idempotents and nilpotents, although it does not immediately give the finer decomposition of the nilpotents provided by the Jordan canonical form. The structure of the nilpotents can easily be obtained from the spectral decomposition. See [8].

Most linear algebra textbooks present the spectral resolution theorem only for special kinds of operators, such as diagonalizable operators. The general case is usually considered as part of the theory of functions of matrices. For this subject the main reference is [6]. See also [4], [7], and [12].

Lancaster and Tismenetsky [7, Ch. 9] use the Jordan canonical form to prove the properties of the component matrices. Hille [5] uses determinants to express the resolvent, and then finds a partial decomposition of the resolvent in which the numerators are the component matrices. Dunford and Schwartz [1, Ch. VII] present a more analytical approach and use Cauchy's integral representation.

In the present paper we show that the spectral resolution of a square matrix A can be obtained in a simple way if we know a nonzero polynomial $w(z)$ such that $w(A)=0$. The polynomial w need not be the characteristic nor the minimal polynomial of A, but, of course, the minimal polynomial must divide w. We use only polynomial interpolation and properties of the map that sends the polynomial $p(z)$ to the matrix $p(A)$. In particular, determinants, canonical forms, inner products, and integrals are not used. We present a construction of the basic Hermite interpolation polynomials based on [9] and [10]. The explicit expressions for these polynomials are our main tools to obtain the properties of the component matrices. We include in section 2 some basic results and present a restricted form of the spectral resolution, related to Lagrange's interpolation, which is the version that appears most often in the literature. We also try to clarify the relationships among resolvents, partial decomposition, and interpolation.

One of the results that we will use frequently is the relationship between polynomial interpolation and the division algorithm for polynomials that we describe next.

Let $w(z)$ be a monic polynomial of degree $n+1$ with roots $\lambda_0, \lambda_1, \ldots, \lambda_s$, which are pairwise distinct, with multiplicities $m_0, m_1, \ldots, m_s$ respectively. The Hermite interpolation theorem, which we prove in section 4, states that for any given numbers $\alpha_{j,k}$ where $0 \le j \le s$ and $0 \le k \le m_j - 1$, there exists a unique polynomial $v(z)$ of degree at most equal to n, such that $v^{(k)}(\lambda_j) = \alpha_{j,k}$, for $0 \le j \le s$ and $0 \le k \le m_j - 1$.

Let $p(z)$ be a polynomial. By the division algorithm there exist unique polynomials q and r such that $p = wq + r$ and, either $r = 0$ or r has degree at most equal to n. Since each λ_j is a root of qw with multiplicity at least equal to m_j, the equation $p = wq + r$ implies that $p^{(k)}(\lambda_j) = r^{(k)}(\lambda_j)$, for $0 \le j \le s$ and $0 \le k \le m_j - 1$. Therefore the remainder $r(z)$ of the division of p by w is the polynomial of degree at most n that interpolates the values $p^{(k)}(\lambda_j)$. This clearly implies the following proposition.

PROPOSITION 1.1. *Let $w(z)$ be as defined above and let p and u be polynomials. Then we have $p \equiv u \mod w$ if and only if $p^{(k)}(\lambda_j) = u^{(k)}(\lambda_j)$, for $0 \le j \le s$ and $0 \le k \le m_j - 1$.*

2. The resolvent and Lagrange's interpolation

Let $w(z)$ be a monic polynomial of degree $n+1$. Define the difference quotient

$$w[z,t] = \frac{w(z) - w(t)}{z - t}.$$

The polynomial identity

$$z^{k+1} - t^{k+1} = (z-t)\sum_{j=0}^{k} z^j t^{k-j} \tag{2.1}$$

implies that $w[z,t]$ is a symmetric polynomial in z and t, of degree n in each variable. If $w(z) = z^{n+1} + b_1 z^n + b_2 z^{n-1} + \cdots + b_{n+1}$ then a simple reordering of summands yields

$$w[z,t] = \sum_{k=0}^{n} w_k(z)\, t^{n-k}, \tag{2.2}$$

where $w_k(z) = z^k + b_1 z^{k-1} + \cdots + b_k$, for $0 \le k \le n$. These are called the *Horner polynomials* of w. It is clear that they form a basis for the vector space $\mathcal{P}_n$ of all polynomials of degree at most equal to n. This basis is often called the control basis [2].

Proposition 2.1. *Let w be a monic polynomial of degree $n+1$ and let $\{f_0, f_1, \ldots, f_n\}$ be a basis of $\mathcal{P}_n$. There exists a unique basis $\{F_0, F_1, \ldots, F_n\}$ of $\mathcal{P}_n$ such that*

$$w[z,t] = \sum_{k=0}^{n} F_{n-k}(t) f_k(z). \tag{2.3}$$

Furthermore, if f_k has degree k then F_k has degree k, for $0 \le k \le n$.

Proof. Let $C = [c_{k,j}]$ be the nonsingular matrix that satisfies

$$t^{n-k} = \sum_{j=0}^{n} c_{k,j} f_{n-j}(t), \qquad 0 \le k \le n.$$

Substitution in (2.2) and the interchange of the sums yields

$$w[z,t] = \sum_{k=0}^{n} w_k(z) \sum_{j=0}^{n} c_{k,j} f_{n-j}(t) = \sum_{j=0}^{n}\sum_{k=0}^{n} c_{k,j}\, w_k(z) f_{n-j}(t).$$

Define

$$F_j(z) = \sum_{k=0}^{n} c_{k,j}\, w_k(z)\,, \qquad 0 \le j \le n\,.$$

Therefore (2.3) holds. Since C^{T} is nonsingular it is clear that the F_j form a basis for $\mathcal{P}_n$. If C is upper triangular then C^{T} is lower triangular. This proves the last part of the assertion.

We will show next how the difference quotient $w[z,t]$ can be used to construct the resolvent of a matrix.

Let $w(z)$ be a monic polynomial of degree $n+1$ and let A be a square matrix of order N with complex entries that satisfies $w(A) = 0$. The polynomial identity $(t-z)w[t,z] = w(t) - w(z)$ gives us

$$(tI - A)w[tI, A] = w(t)I - w(A) = w(t)I\,. \tag{2.4}$$

Therefore, for any complex number t such that $w(t) \neq 0$, we have

$$(tI - A)^{-1} = \frac{w[tI, A]}{w(t)}\,. \tag{2.5}$$

This construction of the resolvent is quite old and has been rediscovered many times. See [3] and [4].

By Proposition 2.1, for each basis $\{f_j\}$ of $\mathcal{P}_n$ we obtain

$$(tI - A)^{-1} = \sum_{k=0}^{n} \frac{F_k(t)}{w(t)} f_{n-k}(A)\,. \tag{2.6}$$

For example, (2.2) yields

$$(tI - A)^{-1} = \sum_{k=0}^{n} \frac{w_k(t)}{w(t)} A^{n-k}\,, \tag{2.7}$$

and

$$(tI - A)^{-1} = \sum_{k=0}^{n} \frac{t^{n-k}}{w(t)} w_k(A)\,. \tag{2.8}$$

Let us consider another example. Let the roots of w be $\lambda_0, \lambda_1, \ldots, \lambda_n$, not necessarily distinct. Define $N_0(z) = 1$ and

$$N_k(z) = (z-\lambda_0)(z-\lambda_1)\cdots(z-\lambda_{k-1}), \qquad 1 \le k \le n\,. \tag{2.9}$$

These are the *Newton polynomials* associated with the sequence of roots $\lambda_0, \lambda_1, \ldots, \lambda_n$. Let F_k be the Newton polynomials associated with the sequence $\lambda_n, \lambda_{n-1}, \ldots, \lambda_0$. Then, by a simple telescopic summation we have

$$w[z,t] = \sum_{k=0}^{n} F_{n-k}(t) N_k(z),$$

and thus

$$(tI - A)^{-1} = \sum_{k=0}^{n} \frac{1}{N_{k+1}(t)} N_k(A). \tag{2.10}$$

Since the only restriction we have imposed on w is that it be monic and $w(A) = 0$, it is possible that some of the roots of w are not in the spectrum of A. Let us see what happens in such a case. Suppose now that $w(z) = u(z)v(z)$ where $u(A) = 0$. Then, since $w[t,z]$ can be written in the form $w[t,z] = u(z)v[t,z] + v(t)u[t,z]$, we obtain

$$\frac{w[tI, A]}{w(t)} = \frac{u[tI, A]v(t)}{u(t)v(t)}.$$

Therefore the roots of v are removable singularities of $w[tI, A]/w(t)$.

Let us now consider a simple case. Let $w(z) = \prod_{j=0}^{n}(z - \lambda_j)$ where the λ_j are pairwise distinct complex numbers. It is obvious that $w[\lambda_j, \lambda_k] = \delta_{j,k} w'(\lambda_k)$. Define the basic *Lagrange interpolation polynomials* associated with the nodes λ_j by

$$\ell_k(z) = \frac{w[z, \lambda_k]}{w'(\lambda_k)}, \qquad 0 \le k \le n. \tag{2.11}$$

Note that ℓ_k is a polynomial of degree n and $\ell_k(\lambda_j) = \delta_{j,k}$. Therefore

$$p(z) = \sum_{k=0}^{n} p(\lambda_k)\ell_k(z), \qquad p \in \mathcal{P}_n. \tag{2.12}$$

This is Lagrange's interpolation formula.

PROPOSITION 2.2.

(i)
$$1 = \sum_{k=0}^{n} \ell_k(z),$$

(ii)
$$z = \sum_{k=0}^{n} \lambda_k \ell_k(z),$$

(iii)
$$\ell_j \ell_k \equiv \delta_{j,k} \ell_k \mod w.$$

Proof. Parts i) and ii) are cases of (2.12).

It is clear that $\ell_j \ell_k$ is a multiple of w if $j \ne k$. Since $\ell_k^2(\lambda_i) = \delta_{k,i}$ the polynomial that interpolates ℓ_k^2 at the roots of w is ℓ_k, and this is the same as the remainder of the division of ℓ_k^2 by w. This proves part iii).

Since $w[z,t]$ is a polynomial in z of degree n, by Lagrange's interpolation formula we have

$$w[z,t] = \sum_{k=0}^{n} w[t,\lambda_k]\ell_k(z)\,. \tag{2.13}$$

This formula gives us

$$(tI - A)^{-1} = \sum_{k=0}^{n} \frac{w[t,\lambda_k]}{w(t)}\ell_k(A) = \sum_{k=0}^{n} \frac{1}{t-\lambda_k}\ell_k(A)\,. \tag{2.14}$$

Replacing z by A in Proposition 2.2 we obtain immediately the following theorem.

THEOREM 2.3 (Spectral Resolution; simple case).
Let A be an $N \times N$ matrix. Suppose that w is a monic polynomial of degree $n+1$ with pairwise distinct roots $\lambda_0, \lambda_1, \ldots, \lambda_n$, such that $w(A) = 0$. Let $\ell_k(z)$ be the basic Lagrange polynomials associated with the λ_k and let $E_k = \ell_k(A)$, for $0 \le k \le n$. Then

$$I = \sum_{k=0}^{n} E_k\,, \tag{i}$$

$$A = \sum_{k=0}^{n} \lambda_k E_k\,, \tag{ii}$$

$$E_j E_k = \delta_{j,k} E_k\,. \tag{iii}$$

Suppose that $w(z) = u(z)v(z)$ and $u(A) = 0$. Then

$$\frac{w[t,z]}{w'(t)} = \frac{u(z)v[t,z] + v(t)u[t,z]}{u(t)v'(t) + u'(t)v(t)}\,,$$

and hence

$$E_j = \ell_j(A) = \frac{v(\lambda_j)u[\lambda_j I, A]}{u(\lambda_j)v'(\lambda_j) + u'(\lambda_j)v(\lambda_j)}\,.$$

Therefore, if λ_j is not in the spectrum of A, that is, if $v(\lambda_j) = 0$ (and thus $u(\lambda_j) \neq 0$) then $E_j = 0$. In the other case we have $u(\lambda_j) = 0$ and $v(\lambda_j)) \neq 0$ and thus $E_j = u[\lambda_j I, A]/u'(\lambda_j)$. This means that we can reduce w to the minimal polynomial of A, and therefore that Theorem 2.3 holds for any diagonalizable matrix. In some textbooks the existence of a spectral decomposition like that of Theorem 2.3 is presented as a condition equivalent to diagonalizability of A.

The case when w is the characteristic polynomial of A is particularly simple, since A is then a matrix of order $n+1$ that has $n+1$ distinct characteristic roots λ_k. Consequently, the image of each idempotent E_k is a one-dimensional subspace and hence E_k is a matrix of rank one. Let V be a matrix such that its k-th column v_k is an eigenvector corresponding to λ_k. Since the v_k are linearly independent, the matrix V is nonsingular. Let x_k be the k-th row of V^{-1}. It is easy to see that $E_k = v_k x_k$. Note that this simple construction of the idempotents E_k does not work if the minimal polynomial has distinct roots but is not equal to the characteristic polynomial of A, since then some of the E_k are projections on subspaces of dimension greater than one.

3. Hermite's interpolation

Let

$$w(z) = \prod_{j=0}^{s} (z - \lambda_j)^{m_j}, \tag{3.1}$$

where the λ_j are distinct and the multiplicities m_j are positive integers with $\sum_j m_j = n+1$. Define the index set

$$\mathcal{I} = \{(j,k) : 0 \le j \le s,\ 0 \le k < m_j\}.$$

Note that $\mathcal{I}$ has $n+1$ elements.

Define the polynomials

$$q_{j,k}(z) = \frac{w(z)}{(z-\lambda_j)^{m_j - k}}, \qquad (j,k) \in \mathcal{I}. \tag{3.2}$$

Note that λ_j is a root of $q_{j,k}$ of multiplicity k, for $k \ge 1$, and is not a root of $q_{j,0}$. Note also that $q_{j,k}(z) = (z-\lambda_j)^k q_{j,0}(z)$. The *Taylor functionals* $T_{j,k}$ are defined by

$$T_{j,k} f = \frac{1}{k!} f^{(k)}(\lambda_j), \qquad (j,k) \in \mathcal{I},$$

for any function f sufficiently differentiable at λ_j. We define the functionals $L_{j,k}$ on the space of polynomials by

$$L_{j,k} p = T_{j,k} \frac{p(z)}{q_{j,0}(z)}, \qquad (j,k) \in \mathcal{I}. \tag{3.3}$$

By Leibniz's rule we have

$$L_{j,k}p = \sum_{i=0}^{k} T_{j,k-i} \frac{1}{q_{j,0}} \ T_{j,i}p\,,$$

and hence $L_{j,k}$ is a linear combination of Taylor functionals.

PROPOSITION 3.1.

$$L_{i,r} q_{j,k} = \delta_{(i,r),(j,k)}, \qquad (i,r),\ (j,k) \in \mathcal{I}\,, \tag{3.4}$$

and hence $\{q_{j,k}\}$ *is a basis of* $\mathcal{P}_n$ *and* $\{L_{i,r}\}$ *is its dual basis.*

The proof is a direct application of Leibniz's rule. See [9].

COROLLARY 3.2 (Lagrange-Sylvester interpolation formula).

$$p(z) = \sum_{(j,k)\in\mathcal{I}} L_{j,k}p \ q_{j,k}(z), \qquad p \in \mathcal{P}_n\,. \tag{3.5}$$

Dividing both sides of (3.5) by $w(z)$ and using (3.2) we obtain the *partial fraction decomposition* formula.

COROLLARY 3.3.

$$\frac{p(z)}{w(z)} = \sum_{(j,k)\in\mathcal{I}} \frac{L_{j,k}p}{(z-\lambda_j)^{m_j-k}}\,, \qquad p \in \mathcal{P}_n\,. \tag{3.6}$$

Leibniz's rule yields

$$T_{j,m_j-1-k} w[z,t] = q_{j,k}(t)\,, \qquad (j,k) \in \mathcal{I}\,, \tag{3.7}$$

where the functional acts with respect to z. We define the polynomials

$$H_{j,k}(t) = L_{j,m_j-1-k} w[z,t]\,, \qquad (j,k) \in \mathcal{I}\,, \tag{3.8}$$

where the functional acts with respect to z. Then, using Leibniz's rule, the definition of the functionals $L_{j,k}$, and (3.7) we get

$$\begin{aligned} H_{j,k}(t) &= \sum_{i=0}^{m_j-1-k} T_{j,i} \frac{1}{q_{j,0}} \ T_{j,m_j-1-k-i} w[z,t] \\ &= \sum_{i=0}^{m_j-1-k} T_{j,i} \frac{1}{q_{j,0}} \ q_{j,k+i}(t)\,. \end{aligned}$$

Therefore

(3.9)
$$H_{j,k}(t) = q_{j,k}(t) \sum_{i=0}^{m_j-1-k} T_{j,i}\frac{1}{q_{j,0}} \, (t-\lambda_j)^i \, .$$

Note that each $H_{j,k}$ is a polynomial of degree n.

By the Lagrange-Sylvester interpolation formula we have

(3.10)
$$w[z,t] = \sum_{(j,k)\in\mathcal{I}} H_{j,k}(z)\, q_{j,m_j-1-k}(t)\,,$$

and thus

(3.11)
$$\frac{w[z,t]}{w(t)} = \sum_{(j,k)\in\mathcal{I}} \frac{H_{j,k}(z)}{(t-\lambda_j)^{k+1}}\,.$$

PROPOSITION 3.4.

(3.12)
$$T_{i,r}H_{j,k} = \delta_{(i,r),(j,k)}\,, \qquad (i,r),\ (j,k)\in\mathcal{I}.$$

Proof. Using the definition of the $H_{j,k}$ and interchanging the linear functionals we get

$$T_{i,r}H_{j,k}(t) = L_{j,m_j-1-k}T_{i,r}w[z,t] = L_{j,m_j-1-k}q_{i,m_i-1-r}(z) = \delta_{(j,k),(i,r)}\,,$$

where $T_{i,r}$ acts with respect to t.

The polynomials $H_{j,k}$ are the basic *Hermite interpolation polynomials* associated with the roots of w.

We say that a function f is defined on the roots of w if $T_{j,k}f$ is defined for $(j,k)\in\mathcal{I}$. Proposition 3.4 gives us immediately the following

PROPOSITION 3.5 (Hermite's interpolation formula).
For any function f defined on the roots of w, the polynomial

(3.13)
$$p(t) = \sum_{(j,k)\in\mathcal{I}} T_{j,k}f\ H_{j,k}(t)$$

is the unique element of $\mathcal{P}_n$ that satisfies $T_{j,k}f = T_{j,k}p$ for $(j,k)\in\mathcal{I}$.

We can write (3.9) in the form

$$H_{j,k}(t) = (t-\lambda_j)^k q_{j,0}(t) \sum_{i=0}^{m_j-1-k} T_{j,i}\frac{1}{q_{j,0}}\, (t-\lambda_j)^i\,.$$

The sum above is the Taylor series of $1/q_{j,0}(t)$ at $t = \lambda_j$, truncated at the $(m_j - 1 - k)$-th power of $(t - \lambda_j)$. A simple computation yields

$$H_{j,k}(t) = (t - \lambda_j)^k - (t - \lambda_j)^{m_j} \sum_{r=0}^{n-m_j} c_r (t - \lambda_j)^r , \tag{3.14}$$

where

$$c_r = \sum_{i=0}^{r} T_{j,i} q_{j,0} \; T_{j,m_j+r-i} \frac{1}{q_{j,0}} . \tag{3.15}$$

PROPOSITION 3.6. *The basic Hermite interpolation polynomials satisfy:*
i) $H_{j,k} H_{i,r} \equiv 0 \mod w$ *if* $j \neq i$,
ii) $H_{j,k} H_{j,r} \equiv (t - \lambda_j)^r H_{j,k}(t) \equiv \begin{cases} H_{j,k+r} \mod w & \text{if } 0 \leq k + r < m_j , \\ 0 \mod w & \text{if } k + r \geq m_j . \end{cases}$

Proof. From the definition of the polynomials $q_{i,k}$ it is clear that $q_{i,r} q_{j,k}$ is a multiple of w if $i \neq j$. From (3.9) we see that $q_{i,r}$ divides $H_{i,r}$ and $q_{j,k}$ divides $H_{j,k}$. Therefore $H_{j,k} H_{i,r}$ is also a multiple of w. This proves part i).

By Proposition 1.1, for any pair of polynomials p and u we have $p \equiv u \mod w$ if and only if $T_{j,k} p = T_{j,k} u$ for $(j,k) \in \mathcal{I}$. Then it is clear that part ii) follows from (3.14), which gives an explicit formula for the expansion of $H_{j,k}$ in powers of $(t - \lambda_j)$.

We will use in the next section the following special cases of Hermite's interpolation formula:

$$1 = \sum_{j=0}^{s} H_{j,0}(z) , \tag{3.16}$$

$$z = \sum_{j=0}^{s} \{ \lambda_j H_{j,0}(z) + H_{j,1}(z) \} . \tag{3.17}$$

The difference quotient $w[z,t]$ can be considered as the kernel function of an interpolation operator, as we show next. Let us define the linear functional Δ_w, called the *divided difference* with respect to the roots of w, as follows. For any function f defined on the roots of w,

$$\Delta_w f = \sum_{j=0}^{s} L_{j,m_j-1} f . \tag{3.18}$$

Since each $L_{j,k}$ is a linear combination of Taylor functionals, so is Δ_w. It is easy to see that

$$\Delta_w f = \sum_{j=0}^{s} \text{Residue of } \frac{f}{w} \text{ at } \lambda_j .$$

Using Proposition 3.1, equation (3.10), and Hermite's interpolation the proof of the following theorem is a simple computation.

THEOREM 3.7 (General interpolation formula).
For any function f defined on the roots of w,

$$p(t) = \Delta_w \{w[z,t]f(z)\} \tag{3.19}$$

is the polynomial of degree at most n that interpolates f at the roots of w.

Note that by the above theorem and Proposition 2.1 we can express the interpolating polynomial in terms of any given basis of the space $\mathcal{P}_n$. See [9] and [10].

4. SPECTRAL RESOLUTION

Let $w(z)$ be as in the previous section a monic polynomial of degree $n+1$ with roots $\lambda_0, \lambda_1, \ldots, \lambda_s$ with multiplicities $m_0, m_1, \ldots, m_s$, respectively. Let A be a square matrix such that $w(A) = 0$. Define

$$E_i = H_{i,0}(A), \quad \text{and} \quad N_i = H_{i,1}(A), \qquad 0 \le i \le s, \tag{4.1}$$

where the $H_{i,k}$ are the basic Hermite interpolation polynomials associated with the roots of w. From Proposition 3.6 and equations (3.16) and (3.17) we obtain immediately the following theorem.

THEOREM 4.1 (Spectral resolution).

i) $A = \sum_{i=0}^{s} \{\lambda_i E_i + N_i\}$,
ii) $I = \sum_{i=0}^{s} E_i$,
iii) $E_i E_j = \delta_{i,j} E_i$,
iv) $N_i = (A - \lambda_i I)E_i = E_i(A - \lambda_i I)$,
v) $N_j E_i = E_i N_j = \delta_{i,j} N_i$,
vi) $N_i N_j = \delta_{i,j} N_i^2$,
vii) *For* $1 \le r \le m_i - 1$ *we have* $N_i^r = H_{i,r}(A) = (A - \lambda_i I)^r E_i$,
viii) $N_i^{m_i} = 0$.

From (3.11) we obtain the following expression for the resolvent of A:

$$(4.2)\qquad (tI - A)^{-1} = \sum_{i=0}^{s} \left\{ \frac{E_i}{(t-\lambda_i)} + \sum_{k=1}^{m_i-1} \frac{N_i^k}{(t-\lambda_i)^{k+1}} \right\}.$$

Note that we can also write this in the form

$$(4.3)\qquad (tI - A)^{-1} = \sum_{i=0}^{s} \left\{ \sum_{k=0}^{m_i-1} \frac{(A-\lambda_i I)^k}{(t-\lambda_i)^{k+1}} \right\} E_i .$$

Suppose now that w has a factorization $w = uv$, where $u(A) = 0$,

$$u(z) = \prod_{j=0}^{r} (z-\lambda_j)^{m_j}, \quad \text{and} \quad v(z) = \prod_{j=r+1}^{s} (z-\lambda_j)^{m_j}.$$

If $j > r$ then $q_{j,0}(z) = u(z)v_{j,0}(z)$, where $v_{j,0}(z) = v(z)/(z-\lambda_j)^{m_j}$. By equation (3.9) the polynomial $q_{j,0}$ is a factor of $H_{j,k}$ and hence u is also a factor of $H_{j,k}$. Therefore $E_j = H_{j,0}(A) = 0$ and $N_j = H_{j,1}(A) = 0$.

If $0 \le j \le r$ then $q_{j,0}(z) = v(z)u_{j,0}(z)$, where $u_{j,0}(z) = u(z)/(z-\lambda_j)^{m_j}$. Then, by (3.8) and (3.3) we have

$$H_{j,k}(t) = T_{j,m_j-i-k} \left\{ \frac{w[t,z]}{q_{j,0}(z)} \right\},$$

and thus

$$\begin{aligned} H_{j,k}(t) &= T_{j,m_j-i-k} \left\{ \frac{u(t)v[t,z] + v(z)u[t,z]}{v(z)u_{j,0}(z)} \right\} \\ &= u(t)T_{j,m_j-i-k} \left\{ \frac{v[t,z]}{v(z)u_{j,0}(z)} \right\} + T_{j,m_j-i-k} \left\{ \frac{u[t,z]}{u_{j,0}(z)} \right\}. \end{aligned}$$

The last term is the basic Hermite interpolation polynomial associated with the roots of $u(z)$, with indices j, k. Let us denote it by $G_{j,k}(t)$. Therefore we have $H_{j,k} \equiv G_{j,k} \mod u$, and consequently $H_{j,k}(A) = G_{j,k}(A)$, since $u(A) = 0$. This means that the roots of w that are not in the spectrum of A do not contribute to the spectral decomposition of A.

We consider next the possibility of reducing the multiplicity of a root λ_j of $w(z)$ for the construction of the spectral resolution of A.

PROPOSITION 4.2. *Suppose that there is an index j such that $E_j \neq 0$ and $N_j^r = 0$ for some r with $1 \le r < m_j$. Let $u(z) = w(z)/(z-\lambda_j)^{m_j-r}$. Then $u(A) = 0$.*

Proof. Since $u(z) = (z - \lambda_j)^r q_{j,0}(z)$ it is clear that $T_{i,k}u$ may be nonzero only if $i = j$ and $k \geq r$. Then, by the Hermite interpolation formula $u(z)$ is a linear combination of the polynomials $H_{j,r}, H_{j,r+1}, \ldots, H_{j,m_j-1}$. By hypothesis $N_j^r = H_{j,r}(A) = 0$ and thus by part vii) of Theorem 4.1, $H_{j,r+i}(A) = 0$ for $i \geq 0$. Therefore $u(A) = 0$.

COROLLARY 4.3. *If w is the minimal polynomial of A then m_j is the index of nilpotency of N_j, for $0 \leq j \leq s$.*

From (3.14) we obtain

$$E_i = I - \sum_{r=0}^{n-m_i} c_r (A - \lambda_i I)^{m_i + r}, \tag{4.4}$$

where the coefficients c_r are given by equation (3.15). From (3.9) we also get

$$E_i = q_{i,0}(A) \sum_{j=0}^{m_i - 1} T_{i,j} \left\{ \frac{1}{q_{i,0}} \right\} (A - \lambda_i I)^j. \tag{4.5}$$

Note that E_i is a polynomial in A of degree n. We show next that the idempotents E_i are essentially unique.

PROPOSITION 4.4. *Let h be an element of $\mathcal{P}_n$ such that $h^2 \equiv h \mod w$. Then $h(z) = \sum_{j=0}^{s} d_j H_{j,0}(z)$ where each d_j is an element of $\{0, 1, -1\}$.*

Proof. The hypothesis $h^2 \equiv h \mod w$ is equivalent to the condition $T_{j,k}h^2 = T_{j,k}h$, for each (j,k) in $\mathcal{I}$. By Leibniz's rule, for each j we must have

$$\sum_{i=0}^{k} T_{j,i}h \; T_{j,k-i}h = T_{j,k}h, \qquad 0 \leq k \leq m_j - 1.$$

This system of equations has only the solutions $T_{j,0}h \in \{0, 1, -1\}$ and $T_{j,k}h = 0$ for $1 \leq k \leq m_j - 1$. Applying the Hermite interpolation formula to h we get the desired conclusion.

COROLLARY 4.5. *Let h be a polynomial such that $h(A)$ is an idempotent. Then*

$$h(A) = \sum_{j=0}^{s} d_j E_j,$$

where the coefficients d_j are elements of $\{0, 1, -1\}$.

The nilpotent matrices N_j have a similar property. Let w be the minimal polynomial of A. Suppose that g is an element of $\mathcal{P}_n$ such that $N = g(A)$ satisfies $N^r = 0$ for some $r \geq 0$. Then w divides g^r and thus $(z - \lambda_j)^{m_j}$ divides g^r for $0 \leq j \leq s$. Therefore $(z - \lambda_j)$ divides g for each j and, by Hermite's interpolation, $g(z)$ is a linear combination of the polynomials $H_{j,k}$ with $k \geq 1$. This means that $N = g(A)$ is a linear combination of the nilpotents N_j^k, with $k \geq 1$.

The spectral resolution of a matrix A is often used to find functions of A. For example, using the properties of the matrices E_i and N_i and the binomial formula we obtain

$$A^r = \sum_{i=0}^{s} \sum_{k=0}^{m_i-1} \binom{r}{k} \lambda_i^{r-k} N_i^k \, E_i \,, \qquad r \geq 0 \,. \tag{4.6}$$

The same formula is obtained by finding the polynomial p that interpolates z^r at the roots of w and then computing $p(A)$, which is

$$p(A) = \sum_{i=0}^{s} \sum_{k=0}^{m_i-1} T_{i,k} z^r H_{i,k}(A) \,.$$

Formula (4.2) for the resolvent of A is obtained in the same way using the polynomial that interpolates $1/(t - z)$, as a function of z, at the roots of w.

The general interpolation formula of Theorem 3.7 yields

$$g(A) = \Delta_w \{ w[zI, A] g(z) \} \,, \tag{4.7}$$

for any function g defined on the roots of w. For example, for $g(z) = e^{tz}$ we get

$$e^{tA} = \sum_{i=0}^{s} \sum_{k=0}^{m_i-1} \frac{t^k}{k!} e^{\lambda_i t} N_i^k E_i \,. \tag{4.8}$$

Using formula (2.2) for $w[z, x]$ we get

$$e^{tA} = \sum_{k=0}^{n} f_k(t) A^{n-k} \,, \tag{4.9}$$

where $f_k(t) = \Delta_w \{ e^{tz} w_k(z) \}$ and the divided difference functional acts with respect to the variable z. See [11] and [12] for some related formulas and applications to the solution of matrix differential equations.

Let us note that (4.7) can be written in the form

$$g(A) = \sum_{i=0}^{s} \text{ Residue of } \left\{ g(z) \frac{w[zI, A]}{w(z)} \right\} \text{ at } \lambda_i \,.$$

Since $w[zI, A]/w(z)$ is the resolvent of A, this formula is analogous to the Cauchy integral representation

$$g(A) = \frac{1}{2\pi i} \int_C g(z)(zI - A)^{-1} \, dz \,,$$

where C is a simple curve whose interior contains the λ_j. See [1, Ch. VII].

REFERENCES

[1] DUNFORD, N. and J. T. SCHWARTZ. *Linear Operators.* Part I. Wiley-Interscience, 1967.

[2] FUHRMANN, P. A. *A Polynomial Approach to Linear Algebra.* Springer-Verlag, 1996.

[3] FRAZER, R. A., W. J. DUNCAN and A. R. COLLAR. *Elementary Matrices.* Cambridge Univ. Press, 1946.

[4] GANTMACHER, F. R. *The Theory of Matrices.* Vol. I. Chelsea, 1960.

[5] HILLE, E. *Methods in Classical and Functional Analysis.* Addison-Wesley, 1972.

[6] HORN, R. A. and C. R. JOHNSON. *Topics in Matrix Analysis.* Cambridge Univ. Press, 1991.

[7] LANCASTER, P. and M. TISMENETSKY. *The Theory of Matrices.* 2nd. ed. Academic Press, 1985.

[8] SOBCZYK, G. The generalized spectral decomposition of a linear operator. *College Math. J. 28* (1997), 27–38.

[9] VERDE-STAR, L. Biorthonormal systems, partial fractions, and Hermite interpolation. *Adv. in Appl. Math. 10* (1989), 348–357.

[10] —— Biorthogonal polynomial bases and Vandermonde-like matrices. *Stud. Appl. Math. 95* (1995), 269–295.

[11] —— Operator identities and the solution of linear matrix difference and differential equations. *Stud. Appl. Math. 91* (1994), 153–177.

[12] —— Functions of matrices. *Linear Algebra Appl. 406* (2005), 285–300.

(Reçu le 28 février 2006)

Luis Verde-Star

Departamento de Matemáticas
UAM-Iztapalapa
Apartado 55–534
México D.F. 09340
México
e-mail: verde@xanum.uam.mx

Leere Seite
Blank page
Page vide