

## On the Computation of AN

Saber N. Elaydi; William A. Harris, Jr.

SIAM Review, Vol. 40, No. 4. (Dec., 1998), pp. 965-971.

Stable URL:

http://links.jstor.org/sici?sici=0036-1445%28199812%2940%3A4%3C965%3AOTCOA%3E2.0.CO%3B2-4

SIAM Review is currently published by Society for Industrial and Applied Mathematics.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <a href="http://www.jstor.org/about/terms.html">http://www.jstor.org/about/terms.html</a>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <a href="http://www.jstor.org/journals/siam.html">http://www.jstor.org/journals/siam.html</a>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.

## ON THE COMPUTATION OF $A^{N*}$

SABER N. ELAYDI<sup>†</sup> AND WILLIAM A. HARRIS, JR.<sup>‡</sup>

This paper is dedicated to Professor Hugh L. Turrittin on his 90th birthday.

**Abstract.** Methods, which are based on the Cayley–Hamilton theorem, for the computation of  $A^n$  for nonsingular A are presented.

Key words. Cayley-Hamilton, difference equation, Casorati matrix

AMS subject classifications. 15-01, 15A21, 39-01, 39A10

PII. S0036144597319235

Let A be a  $k \times k$  nonsingular matrix of constants; then a fundamental matrix for the linear system of difference equations x(n+1) = Ax(n) is  $X(n) = A^n$ .

$$X(n+1) = AX(n), \quad X(0) = I,$$

and the solution of the difference initial value problem

$$x(n+1) = Ax(n) + b(n), \quad x(0) = x_0,$$

is given by

$$x(n) = A^{n}x_{0} + \sum_{j=0}^{n-1} A^{n-j-1}b(j)$$
$$= X(n)x_{0} + X(n)\sum_{j=0}^{n-1} X^{-1}(j+1)b(j)$$

which is analogous to

$$x(t) = e^{At}x_0 + e^{At} \int_{0}^{t} e^{-As}b(s)ds$$

as the solution of the corresponding differential initial value problem. Hence, for linear systems of difference equations, the computation of  $A^n$  is the analogous problem to the computation of  $e^{At}$  for linear systems of differential equations.

One could use transformation methods, writing  $A^n = P(P^{-1}AP)^n P^{-1}$  for non-singular P to compute  $(P^{-1}AP)^n$  and hence  $A^n$  by using Schur's canonical form (triangular) or the Jordan canonical form, e.g., for  $J = (\lambda I + N)$ , with  $N^s = 0$  we have  $J^n = \lambda^n I + \binom{n}{1} \lambda^{n-1} N + \cdots + \binom{n}{s-1} \lambda^{n-s+1} N^{s-1}$ . Such transformation methods require the determination of eigenvectors and generalized eigenvectors and can be computationally involved and tedious.

We present in this note alternative methods for the computation of  $A^n$  for non-singular A which are pedagogically simpler, based on the Cayley–Hamilton theorem

<sup>\*</sup>Received by the editors March 31, 1997; accepted for publication December 12, 1997.

http://www.siam.org/journals/sirev/40-4/31923.html

<sup>&</sup>lt;sup>†</sup>Department of Mathematics, Trinity University, San Antonio, TX 78212-7200 (selaydi@trinity.edu).

<sup>&</sup>lt;sup>‡</sup>Department of Mathematics, University of Southern California, Los Angeles, CA 90089-1113 (wharris@math.usc.edu).

and the exact determination of the eigenvalues of A, which are analogous to methods of Putzer [9], Fulmer [2], Kirchner [6], Hirsch and Smale [3], Leonard [8], and Hsieh, Kohno, and Sibuya [4] for computing  $e^{At}$ .

Let

(1) 
$$c(\lambda) = \det(\lambda I - A) = \lambda^k + c_{k-1}\lambda^{k-1} + \dots + c_1\lambda + c_0$$
$$= (\lambda - \lambda_1)(\lambda - \lambda_2) \dots (\lambda - \lambda_k)$$

be the characteristic polynomial for A. Then the Cayley–Hamilton theorem states that c(A) = 0, or

$$A^{k} + c_{k-1}A^{k-1} + \dots + c_{1}A + c_{0}I = 0$$

and

$$(A - \lambda_1 I)(A - \lambda_2 I) \cdots (A - \lambda_k I) = 0.$$

THEOREM 1. Let A be a  $k \times k$  nonsingular matrix with eigenvalues  $\lambda_1, \ldots, \lambda_k$ , and let  $M(0) = I, M(j) = \prod_{i=1}^{j} (A - \lambda_i I), j \geq 1$ . Then, if  $u_j(n)$  satisfies the (recursive) linear difference system

$$u_1(n+1) = \lambda_1 u_1(n),$$
  $u_1(0) = 1,$   $u_{j+1}(n+1) = \lambda_{j+1} u_{j+1}(n) + u_j(n),$   $u_{j+1}(0) = 0, j = 1, \dots, k-1,$ 

for  $n \geq k$ ,

$$A^{n} = \sum_{j=0}^{k-1} u_{j+1}(n)M(j).$$

We note that  $u_1(n) = \lambda_1^n$  and  $u_{j+1}(n) = \sum_{i=0}^{n-1} \lambda_{j+1}^{n-i-1} u_j(i), j = 1, \dots, k-1$ , and  $M(k) = \prod_{i=1}^k (A - \lambda_i I) = 0$  (Cayley–Hamilton). Since M(j) is the monic polynomial  $M(j) = A^j$  + lower order terms,  $A^n, n \geq k$ , can also be written as a polynomial in A of degree at most k-1.

THEOREM 2. Let A be a  $k \times k$  nonsingular matrix with characteristic polynomial  $c(\lambda) = \lambda^k + c_{k-1}\lambda^{k-1} + \cdots + c_1\lambda + c_0$ , and let z(n) be the solution of the scalar kth order difference equation.

$$z(n+k) + c_{k-1}z(n+k-1) + \cdots + c_1z(n+1) + c_0z(n) = 0,$$

$$z(0) = z(1) = \cdots = z(k-2) = 0, \quad z(k-1) = 1,$$

and

$$\begin{pmatrix} q_1(n) \\ \vdots \\ c_1 \\ \vdots \\ q_k(n) \end{pmatrix} = \begin{pmatrix} c_1 & c_2 & \cdot & \cdot & \cdot & c_{k-1} & 1 \\ c_2 & \cdot & \cdot & \cdot & \cdot & 1 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ c_{k-1} & 1 & \cdot & & & \ddots & \vdots \\ 1 & 0 & \cdot & \cdot & \cdot & 0 \end{pmatrix} \begin{pmatrix} z(n) \\ \vdots \\ \vdots \\ z(n+k-1) \end{pmatrix}.$$

Then, for  $n \geq k$ ,

$$A^{n} = \sum_{j=0}^{k-1} q_{j+1}(n)A^{j}.$$

We note that the difference equation  $z(n+k)+c_{k-1}z(n+k-1)+\cdots+c_0z(n)=0$  can conveniently be written as c(E)z(n)=0 using the shift operator Ez(n)=z(n+1). Further, if the matrix A is a companion matrix, the linear difference system x(n+1)=Ax(n) is equivalent to the kth order scalar difference equation c(E)z(n)=0, and it is not surprising that solutions of c(E)z(n)=0 would play a role in computing  $A^n$ . That this is true in general follows from the observation

$$E(A^n) = A^{n+1} = A(A^n)$$

and hence if  $p(\lambda)$  is any polynomial,  $p(E)(A^n) = p(A)A^n$ , and thus for any annihilating polynomial  $p(\lambda)$  for A,  $p(E)A^n = 0$ . Since the Cayley-Hamilton theorem simply states that  $c(\lambda)$  is an annihilating polynomial for A, every element of  $A^n$  satisfies the kth order scalar difference equation c(E)y(n) = 0. The minimal polynomial,  $m(\lambda)$ , is also an annihilating polynomial for A, but is easily determined only in special cases, e.g., when A is a companion matrix  $(m(\lambda) = c(\lambda))$  or A real symmetric,  $m(\lambda) = \prod_{i=1}^{s} (\lambda - \mu_j)$ , where  $\mu_j$  are the distinct eigenvalues of A.

THEOREM 3. Let A be a  $k \times k$  nonsingular matrix with characteristic polynomial  $c(\lambda) = \lambda^k + c_{k-1}\lambda^{k+1} + \cdots + c_1\lambda + c_0$ , and let  $\{y_1(n), \dots, y_k(n)\}$  be a linearly independent set of solutions of the kth order scalar difference equation c(E)y(n) = 0. Then there exist constant matrices  $E_1, \dots, E_k$  such that

$$A^n = y_1(n)E_1 + y_2(n)E_2 + \dots + y_k(n)E_k$$
.

We note that if the eigenvalues of A are distinct,  $\{\lambda_1^n, \ldots, \lambda_k^n\}$  is a suitable linearly independent set of solutions for c(E)y(n) = 0, and the general form for  $A^n$ , n = 0, 1 yields

$$I = E_1 + E_2 + \dots + E_k$$
  

$$A = \lambda_1 E_1 + \lambda_2 E_2 + \dots + \lambda_k E_k.$$

Further, it can be shown that  $E_iE_j=E_jE_i=0, i\neq j$ , and hence  $E_i^2=E_i$ . Thus we have a resolution of the identity and the spectral representation of A, and hence

$$A^n = (\lambda_1 E_1 + \dots + \lambda_k E_k)^n = \lambda_1^n E_1 + \dots + \lambda_k^n E_k.$$

The SN decomposition of a matrix is a natural generalization of the spectral representations of A when the eigenvalues of A are not distinct. There exists a unique decomposition of the matrix A, A = S + N, SN = NS with S semisimple (similar to a diagonal matrix) and N nilpotent ( $N^k = 0$ ).

THEOREM 4. Let A be a  $k \times k$  nonsingular matrix, and let A = S + N be the unique SN decomposition, SN = NS, S semisimple, N nilpotent, then

$$A^{n} = (S+N)^{n} = S^{n} + \binom{n}{1} S^{n-1} N + \binom{n}{2} S^{n-2} N^{2} + \dots + \binom{n}{k-1} S^{n-k+1} N^{k-1}.$$

Proof of Theorem 1. Let  $\Phi(n) = \sum_{j=0}^{k-1} u_{j+1}(n) M(j)$ . We need only show that  $\Phi(0) = I$  and  $\Phi(n+1) = A\Phi(n)$ . Clearly,  $\Phi(0) = \sum_{j=1}^{k-1} u_{j+1}(0) M(j) = u_1(0) M(0) = I$ . Using the fact that  $AM(j) = (A - \lambda_{j+1}I + \lambda_{j+1}I) M(j) = M(j+1) + \lambda_{j+1}M(j)$ , we have

$$\Phi(n+1) - A\Phi(n) = \sum_{j=0}^{k-1} u_{j+1}(n+1)M(j) - \sum_{j=0}^{k-1} u_{j+1}(n)AM(j)$$

$$= \sum_{j=0}^{k-1} [u_{j+1}(n+1) - \lambda_{j+1}u_{j+1}(n)]M(j) - \sum_{j=0}^{k-1} u_{j+1}(n)M(j+1).$$

Writing  $\sum_{j=0}^{k-1} u_{j+1}(n)M(j+1) = \sum_{i=1}^k u_i(n)M(i) = \sum_{j=1}^{k-1} u_j(n)M(j)$ , since M(k)=0. Thus

$$\Phi(n+1) - A\Phi(n) = [u_1(n+1) - \lambda_1 u_1(n)] + \sum_{j=1}^{k-1} [u_{j+1}(n+1) - \lambda_{j+1} u_{j+1}(n) - u_j(n)]M(j)$$

and the theorem is proved.  $\Box$ 

Remark. Theorem 1 has appeared in LaSalle [7], Kelley and Peterson [5], and Elaydi [1].

*Proof of Theorem* 2. Let  $\Psi(n) = \sum_{j=0}^{k-1} q_{j+1}(n) A^j$ . Clearly,  $\Psi(0) = \sum_{j=0}^{k-1} q_{j+1}(0) A^j = q_1(0) I = I$ . Now

$$\Psi(n+1) - A\Psi(n) = \sum_{j=0}^{k-1} q_{j+1}(n+1)A^j - \sum_{j=0}^{k-1} q_{j+1}(n)A^{j+1}$$
$$= q_1(n+1)I + \sum_{j=1}^{k-1} [q_{j+1}(n+1) - q_j(n)]A^j - q_k(n)A^k,$$

and using  $A^k + c_k A^{k-1} + \cdots + c_1 A + c_0 I = 0$ ,  $\Psi(n+1) - A\Psi(n) = 0$  requires that

$$\begin{array}{rcl} q_1(n+1) & +c_0q_k(n) & = & 0, \\ q_2(n+1) & +c_1q_k(n) & = & q_1(n), \\ & \cdot & \\ & \cdot & \\ q_k(n+1) & +c_{k-1}q_k(n) & = & q_{k-1}(n). \end{array}$$

Defining  $q_k(n) = z(n)$  and applying  $E^{j-1}$  to the jth equation yields

$$q_1(n+1) + c_0 z(n) = 0,$$
  
 $q_2(n+2) + c_1 z(n+1) = q_1(n+1),$   
 $\vdots$   
 $z(n+k) + c_{k-1} z(n+k-1) = q_{k-1}(n+k-1).$ 

Adding these equations yields c(E)z(n) = 0 and back substitution in the first set of equations completes the proof of the theorem.

Proof of Theorem 3.  $A^n = y_1(n)E_1 + \cdots + y_k(n)E_k$  will be valid if the following system of equations has a unique solution:

The coefficient matrix of this system is the Casorati matrix which is nonsingular for a linearly independent set of solutions  $(y_1(n), \ldots, y_k(n))$  of the kth order scalar difference equations c(E)y(n) = 0, and thus the theorem is proved.

Remark. If the eigenvalues of A are distinct, then  $y_1(n) = \lambda_1^n, \ldots, y_k(n) = \lambda_k^n$  is a suitable linearly independent set of solutions of c(E)y(n) = 0 and the Casorati matrix is a Vandermonde matrix and  $E_i = e_i(A)$ , where  $e_i(\lambda) = \prod_{j \neq i} \frac{\lambda - \lambda_j}{\lambda_i - \lambda_j}$  are the Lagrange interpolating polynomials. We note that the nonzero columns of  $E_j$  are eigenvectors if the  $\lambda_i$  are distinct and are generalized eigenvector chains in the general case.

Proof of Theorem 4. Following Hsieh, Kohno, and Sibuya [4], let  $\mu_j$  be the distinct eigenvalues of the matrix A and write the characteristic polynomial  $c(\lambda) = \prod_{i=1}^{s} (\lambda - \mu_i)^{n_i}$  with integers  $n_i \geq 1, n_1 + \cdots + n_s = k$ . Let the partial fraction expression of  $c(\lambda)^{-1}$  be

$$\frac{1}{c(\lambda)} = \frac{c_1(\lambda)}{(\lambda - \mu_1)^{n_1}} + \frac{c_2(\lambda)}{(\lambda - \mu_2)^{n_2}} + \dots + \frac{c_s(\lambda)}{(\lambda - \mu_s)^{n_s}},$$

and define the polynomials  $f_i(\lambda) = c_i(\lambda) \prod_{j \neq i} (\lambda - \mu_j)^{n_j}$ . Then  $1 = f_1(\lambda) + f_2(\lambda) + \cdots + f_s(\lambda)$ , and with  $F_i = f_i(A)$ , we have the resolution of the identity

$$I = F_1 + F_2 + \dots + F_s$$
,  $F_i F_j = F_j F_i = 0$ ,  $i \neq j$  and  $F_i^2 = F_i$ ,

using the Cayley–Hamilton theorem.

Writing  $S = \sum_{i=1}^{s} \mu_i F_i$ , N = A - S, we have A = S + N, SN = NS,  $N^k = 0$ , S semisimple, and the theorem is proved.

*Remark.* The SN decomposition is essentially in Kirshner [6], and is explicitly stated and proved in Hirsch and Smale [3], but we prefer the construction of Hsieh, Kohno, and Sibuya [4].

*Remark.* The restriction to nonsingular A is natural if we wish  $A^0 = I$ . We note that for  $A \equiv 0$  and

$$A = \left(\begin{array}{cc} 0 & 1 \\ 0 & 0 \end{array}\right),$$

we have  $y(n) \equiv 0$  as a solution of y(n+1) = Ay(n). A preliminary transformation will reduce the general case to the one studied in this note.

Example. Consider x(n+1) = Ax(n) where

$$A = \left(\begin{array}{rrr} 0 & 1 & 1 \\ -2 & 3 & 1 \\ -3 & 1 & 4 \end{array}\right).$$

The characteristic equation  $c(\lambda)$  is

$$c(\lambda) = \lambda^3 - 7\lambda^2 + 16\lambda - 12 = (\lambda - 2)^2(\lambda - 3).$$

$$\bullet A^n = \sum_{j=0}^{k-1} u_{j+1}(n) M(j)$$
, where

$$u_1(n) = 2^n$$
,  $u_2(n) = n2^{n-1}$ ,  $u_3(n) = -2^n + 3^n - n2^{n-1}$ ,

$$M(0) = I,$$
  $M(1) = \begin{pmatrix} -2 & 1 & 1 \\ -2 & 1 & 1 \\ -3 & 1 & 2 \end{pmatrix},$   $M(2) = \begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -2 & 0 & 2 \end{pmatrix}$ , and

$$A^{n} = \begin{pmatrix} 2^{n-1} - 3^{n} - n2^{n-1} & n2^{n-1} & -2^{n} + 3^{n} \\ 2^{n} - 3^{n} - n2^{n-1} & (n+2)2^{n-1} & -2^{n} + 3^{n} \\ 2^{n+1} - 2 \cdot 3^{n} - n2^{n-1} & n2^{n-1} & -2^{n} + 2 \cdot 3^{n} \end{pmatrix}.$$

$$\bullet A^n = \sum_{j=0}^{k-1} q_{j+1}(n) A^j$$
, where

$$A^0 = I, \qquad A^2 = \begin{pmatrix} -5 & 4 & 5 \\ -9 & 8 & 5 \\ -14 & 4 & 14 \end{pmatrix},$$

$$\begin{array}{rcl} q_1(n) & = & -3(1+n)2^n + 4 \cdot 3^n, \\ q_2(n) & = & (8+5n)2^{n-1} - 4 \cdot 3^n, \\ q_3(n) & = & -(2+n)2^{n-1} + 3^n, \end{array}$$

and  $z(n) = q_3(n)$  is a solution of

$$z(n+3) - 7z(n+2) + 16z(n+1) - 12z(n) = 0$$
,  $z(0) = z(1) = 0$ ,  $z(2) = 1$ .

$$\bullet A^n = \sum_{j=1}^k y_j(n) E_j$$
, where  $y_1(n) = 2^n, y_2(n) = n2^{n-1}, y_3(n) = 3^n$ ,

$$E_1 = \begin{pmatrix} 2 & 0 & -1 \\ 1 & 1 & -1 \\ 2 & 0 & -1 \end{pmatrix}, \qquad E_2 = \begin{pmatrix} -1 & 1 & 0 \\ -1 & 1 & 0 \\ -1 & 1 & 0 \end{pmatrix}, \qquad E_3 = \begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -2 & 0 & 2 \end{pmatrix}.$$

$$\bullet A^n = (S+N)^n,$$

$$\frac{1}{c(\lambda)} = \frac{1-\lambda}{(\lambda-2)^2} + \frac{1}{\lambda-3},$$

$$1 = (1 - \lambda)(\lambda - 3) + (\lambda - 2)^2 = f_1(\lambda) + f_2(\lambda), \quad F_i = f_i(A),$$

$$F_1 = \left( egin{array}{ccc} 2 & 0 & -1 \ 1 & 1 & -1 \ 2 & 0 & -1 \end{array} 
ight), \qquad F_2 = \left( egin{array}{ccc} -1 & 0 & 1 \ -1 & 0 & 1 \ -2 & 0 & 2 \end{array} 
ight)$$

$$S = 2F_1 + 3F_2 = \left( egin{array}{ccc} 1 & 0 & 1 \ -1 & 2 & 1 \ -2 & 0 & 4 \end{array} 
ight), \qquad N = A - S = \left( egin{array}{ccc} -1 & 1 & 0 \ -1 & 1 & 0 \ -1 & 1 & 0 \end{array} 
ight).$$

$$N^2 = 0$$
,  $SN = NS = 2N$ ,  $A^n = S^n + n2^{n-1}N$ .

Examination of  $F_1$  and  $F_2$  shows that

$$\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

are eigenvectors of S corresponding to the eigenvalue 2, and

$$\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$$

is an eigenvector of S corresponding to the eigenvalue 3.

Thus for

$$P = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 2 \end{pmatrix}, \qquad P^{-1}SP = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

and 
$$S^n = P(P^{-1}SP)^n P^{-1} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 2 \end{pmatrix} \begin{pmatrix} 2^n & 0 & 0 \\ 0 & 2^n & 0 \\ 0 & 0 & 3^n \end{pmatrix} \begin{pmatrix} 2 & 0 & -1 \\ 1 & 1 & -1 \\ -1 & 0 & 1 \end{pmatrix}.$$

## REFERENCES

- [1] S. ELAYDI, An Introduction to Difference Equations, Springer-Verlag, Berlin, 1996.
- [2] E.P. Fulmer, Computation of the matrix exponential, Amer. Math. Monthly, 82 (1975), pp. 156-159.
- [3] M.W. HIRSCH AND S. SMALE, Differential Equations, Dynamical Systems and Linear Algebra, Academic Press, New York, 1974.
- [4] P.F. HSIEH, M. KOHNO, AND Y. SIBUYA, Construction of a fundamental matrix solution at a singular point of the first kind by means of the SN decomposition of matrices, Linear Algebra Appl., 239 (1996), pp. 26-76.
- [5] W.G. KELLEY AND A.C. PETERSON, Difference Equations, An Introduction with Applications, Academic Press, San Diego, CA, 1991.
- [6] R.B. KIRCHNER, An explicit formula for  $e^{At}$ , Amer. Math. Monthly, 74 (1967), pp. 1200–1204.
- [7] J.P. LASALLE, Stability theory for difference equations, in MAA Studies in Mathematics 14, Studies in Ordinary Differential Equations, J. Hale, ed., Mathematical Association of America, Washington, DC, 1977, pp. 1–31.
- [8] I.E. LEONARD, The matrix exponential, SIAM Rev., 38 (1996), pp. 507-512.
- E.J. PUTZER, Avoiding the Jordan cononical form in the discussion of linear systems with constant coefficients, Amer. Math. Monthly, 73 (1966), pp. 2-7.