REPRINTED FROM

The Fibonacci Quarterly

VOLUME 28 February 1990 NUMBER 1

SUMS OF POWERS BY MATRIX METHODS

Dan Kalman
The Aerospace Corporation, P.O. Box 92957, Los Angeles, CA 90009-2957
(Submitted March 1988)

Let $\{s_n^r\}$ be the sequence of sums of r^{th} powers given by

$$(1)\qquad s_n^r = \sum_{k=0}^{n} k^r.$$

These familiar sequences are the subject of an extensive literature, a few recent samples of which may be found among the references. The present note has two objectives:

- To illustrate the application of matrix methods in the context of finite difference equations; and
- To publicize the following beautiful matrix formula for s_n^r.

$$(2)\qquad s_n^r = \left[\binom{n}{1}\ \binom{n}{2}\ \binom{n}{3}\ \cdots\ \binom{n}{r+1}\right]\begin{bmatrix} 1 & & & & \\ -1 & 1 & & & \\ 1 & -2 & 1 & & \\ \vdots & \vdots & \vdots & \ddots & \\ \pm 1 & \mp\binom{r}{1} & \pm\binom{r}{2} & \cdots & 1 \end{bmatrix}\begin{bmatrix} 1^r \\ 2^r \\ 3^r \\ \vdots \\ (r+1)^r \end{bmatrix}$$

This equation may be viewed as reducing the sum of n terms of the power sequence to a linear combination of the first $r + 1$ terms. Accordingly, there is an implicit assumption that $n > r$. Note that the matrix appearing as the middle factor on the right side of this equation is lower triangular. The zeros that should appear above the main diagonal have been omitted. The nonzero entries constitute a version of Pascal's triangle with alternating signs.

The scalar equivalent of equation (2) has appeared previously ([4], eq. 57, p. 33) and can be derived by standard elementary manipulations of series expansions for exponential functions. The main virtue of the matrix form is esthetic: it reveals a nice connection between s_n^r and Pascal's triangle, and is easily remembered.

The main idea we wish to present regarding the application of matrix analysis to difference equations may be summarized as follows. In general, an n^{th}-order difference equation with constant coefficients is expressible as a first-order vector equation. The solution of this first-order vector equation is given in terms of powers of the coefficient matrix. By reducing the coefficient matrix to its Jordan canonical form, the powers can be explicitly calculated, finally leading to a formula for a solution to the original difference equation. This approach was discussed previously [5] for the case in which the matrix is diagonalizable. In applying this method to the derivation of (2), the matrix is not diagonalizable. Another example with a nondiagonalizable matrix will also be presented, connected with reference [3].

In the interest of completeness, a few results about linear difference equations will be presented. These can also be found in any introductory text on the subject, for example [8].

The sequence s_n^r satisfies the recursive relation

$$s_{n+1}^r = s_n^r + (n + 1)^r.$$

This is an example of a difference equation. It expresses one term of a sequence as a function of the preceding term, and the sequence index n. More generally, a difference equation of order k specifies a term of a sequence $\{a_n\}$ as a function of the preceding k terms and n. We shall be especially concerned with linear, constant coefficient, homogeneous difference equations. Any equation of this type can be cast in the form

$$(3) \qquad a_{n+k} + c_{k-1}a_{n+k-1} + \cdots + c_0 a_n = 0; \quad n \geq 0,$$

where the coefficients c_j are constants. Hereafter, we assume all difference equations are of this type. Clearly, given initial terms a_0 through a_{k-1}, the remaining terms of the sequence are uniquely determined by equation (3). The main objective of the next section is to develop techniques to express these terms as a function of n.

The analysis of difference equations is expedited by reformulating equation (3) in terms of linear operators. Accordingly, we focus for the present on the linear space of sequences $\{a_k\}_{k=0}^{\infty}$ of complex numbers, and state

Definition 1: The linear operator L, called the *lag operator*, is defined by the relation

$$(4) \qquad \mathrm{L}\{a_n\}_{n=0}^{\infty} = \{a_{n+1}\}_{n=0}^{\infty}.$$

L has the effect of *shifting* the terms of a sequence by one position. Thus, it is often convenient to write

$$\mathrm{L}a_n = a_{n+1}.$$

Now (3) may be expressed in the form

$$p(\mathrm{L})\{a_n\} = 0$$

where $p(t) = t^k + c_{k-1}t^{k-1} + \cdots + c_0$ is called the characteristic polynomial of the equation. We follow the usual convention that the constant term of the polynomial p operates on the sequence $\{a_n\}$ by scalar multiplication. Since $p(\mathrm{L})$ is a linear operator, solving (3) amounts to determining the null space.

Now we turn our attention to the application involving s_n^r. As a first step, we use the operator approach to characterize polynomials in n as solutions to a specific class of difference equations. The statement and proof of this result will be simplified by the following notation.

Definition 2: D is the operator L − 1. N_k is the null space of D^k. Note that $\mathrm{N}_1 \subseteq \mathrm{N}_2 \subseteq \mathrm{N}_3 \ldots$.

Theorem 1: For any k, N_k consists of the sequences $\{a_n\}$ such that $a_n = p(n)$ for some polynomial p of degree less than k.

Proof: We show first that polynomials of degree less than k are contained in N_k. For $k = 1$, with a_n a polynomial of degree 0, and thus constant, it is clear that $\mathrm{D}\{a_n\} = 0$. Proceeding by induction, assume that polynomials of degree less than $k - 1$ are in N_{k-1}, and hence in N_k. Showing that $\{n^{k-1}\}$ is in N_k then assures that all polynomials of degree less than or equal to $k - 1$ are contained in N_k. One application of D to $\{n^{k-1}\}$ produces the sequence

$$\{(n + 1)^{k-1} - n^{k-1}\}.$$

This result is a polynomial of degree $k - 2$ and so is annihilated by D^{k-1}, by the induction hypothesis. This shows that $D^k\{n^{k-1}\} = 0$, and completes the first part of the proof.

For the converse, we must show that the polynomials of degree less than k exhaust N_k. Since these polynomials comprise a subspace of dimension k, it will suffice to show that N_k has dimension no more than k. This statement is clearly true for the case that $k = 1$. As before, the general case shall be established by induction.

Assume that N_j has dimension j for all j less than k, and suppose that a_n and b_n are in N_k but not in N_{k-1}. Then $D^{k-1}a_n$ and $D^{k-1}b_n$ are nonzero elements of N_1, which is one dimensional. This implies that, for some scalar c,

$$D^{k-1}a_n = cD^{k-1}b_n.$$

Hence, $a_n - cb_n$ lies in N_{k-1}. We conclude that the dimension of N_k can exceed that of N_{k-1} by at most 1. Finally, by the induction hypothesis, the dimension of N_k is no more than k, completing the proof. □

This result may be immediately applied to the analysis of s_n^r. As observed previously,

$$s_{n+1}^r - s_n^r = (n + 1)^r$$

which is, in operator notation,

$$Ds_n^r = (n + 1)^r.$$

Now the right side is a polynomial in n of degree r, so is annihilated by D^{r+1}. Thus, applying D^{r+1} to both sides yields

$$D^{r+2}s_n^r = 0$$

and hence, s_n^r is in N_{r+2}. Moreover,

$$D^{r+1}s_n^r = D^r(n + 1)^r,$$

which is not zero. It follows that

$$s_n^r \in N_{r+2}\backslash N_{r+1},$$

and that s_n^r is a polynomial in n of degree $r + 1$.

The realization of s_n^r as a solution to the equation

$$D^{r+2}a_n = 0$$

is more significant for our purposes than is the characterization of s_n^r as a polynomial. For future reference, it is convenient to express this equation in the form

$$(5) \qquad (L - 1)^{r+2}a_n = 0.$$

We show next how matrix methods can be employed to solve difference equations. Then, as a particular example, we apply the method to (5) to derive (2).

2. Matrix Methods for Difference Equations

Matrices appear as the result of a standard device for transforming a k^{th}-order scalar equation into a linear vector equation. The transformation is perfectly analogous to one used in the analysis of differential equations ([1], p. 192), and was used in the form presented below in [7].

Suppose a_n satisfies a difference equation of order k, as in equation (3). For each $n \geq 0$, define the k-dimensional vector $\mathbf{v}_n$ according to

$$\mathbf{v}_n = \begin{bmatrix} a_n \\ a_{n+1} \\ a_{n+2} \\ \vdots \\ a_{n+k-1} \end{bmatrix}$$

The vector $\mathbf{v}_n$ may be visualized as a window displaying k entries in the infinite column

$$\begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \end{bmatrix}$$

Now the transformation from $\mathbf{v}_n$ to $\mathbf{v}_{n+1}$ can be formulated as multiplication by the $k \cdot k$ matrix C given by

$$\mathrm{C} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -c_0 & -c_1 & -c_2 & \cdots & -c_{k-1} \end{bmatrix}$$

This matrix is the companion matrix for the characteristic polynomial of the original difference equation (3). It can also be understood as a combination of row operations. In this view, C has the effect of rolling rows 2 through k up one position, and creating in place of row k the linear combination

$$-c_0(\text{row } 1) - c_1(\text{row } 2) - \cdots - c_{k-1}(\text{row } k).$$

These operations correspond exactly to the transformation from $\mathbf{v}_n$ to $\mathbf{v}_{n+1}$.

Visually, multiplying $\mathbf{v}_n$ by C has the effect of moving the window described earlier down one position. Algebraically, $\mathbf{v}_n$ satisfies the *vector* difference equation

(6) $\qquad \mathbf{v}_{n+1} = \mathrm{C}\mathbf{v}_n.$

Evidently, a solution $\mathbf{v}_n$ of (6) may be characterized by

(7) $\qquad \mathbf{v}_n = \mathrm{C}^n\mathbf{v}_0,$

and so, the solution a_n of (3) is given as the first component of the right side of (7). These remarks may be summarized by expressing a_n in the equation

(8) $$a_n = [1 \quad 0 \quad 0 \cdots 0] \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -c_0 & -c_1 & -c_2 & \cdots & -c_{k-1} \end{bmatrix}^n \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_{k-1} \end{bmatrix}$$

This formula is not really useful as a functional representation of a_n because the powers of C must be computed by what is essentially a recursive procedure (although the computation can be made very efficient by exploiting the special structure of C, see [2]). However, if the roots of the characteristic polynomial are known, the Jordan canonical form of C can be explicitly formulated as described in [7]. Thus, if the reduction of C to its Jordan form

J is expressed by $C = SJS^{-1}$, then the matrix C^n may be replaced by SJ^nS^{-1} in (8). The special case in which the roots are distinct features a diagonal J and so the powers of J are simply expressed. This case is discussed in [5]. In the sequel, we shall focus on the application of matrix methods to the analysis of s_n^r, based on equation (5). Observe that the characteristic polynomial is given by $(t - 1)^{r+2}$, hence, rather than distinct roots, we have a single root of multiplicity $r + 2$. The next section will discuss the properties of the Jordan form for this case, and derive equation (2).

3. Analysis of s_n^r

As observed in the preceding section, s_n^r satisfies the difference equation (5) with the characteristic polynomial $(t - 1)^{r+2}$, and with k therefore equal to $r + 2$. Using this information, the general equation (8) may be particularized to give

$$(9) \qquad s_n^r = [1 \quad 0 \quad 0 \cdots 0]\,\mathbf{C}^n \begin{bmatrix} s_0^r \\ s_1^r \\ s_2^r \\ \vdots \\ s_{r+1}^r \end{bmatrix}$$

C is the companion matrix for $(t - 1)^{r+2}$. It can be shown that the Jordan canonical form for the companion matrix of a polynomial has one Jordan block for each distinct root. (A simple proof of this assertion may be constructed using Theorems 4.5 and 8.5 of [9].) In the present case, the Jordan form J is therefore a single Jordan block corresponding to the root 1. That is, J is a square matrix of dimension $r + 2$ with entries of 1 along the main diagonal and first superdiagonal, and all other entries zero. It will be convenient to write $J = I + N$, where I is the identity matrix. The matrix N is familiar as a nilpotent matrix whose j^{th} power has 1's on the j^{th} superdiagonal, and 0's elsewhere, for $0 \le j \le r + 1$. Accordingly, J^n may be computed as

$$(I + N)^n = \sum_{j=0}^{r+1} \binom{n}{j} N^j,$$

and we observe that this result has constants along each diagonal. Specifically, it is an upper-triangular matrix with 1's on the main diagonal, $\binom{n}{1}$'s on the super diagonal, $\binom{n}{2}$'s on the next diagonal, and so on.

The matrix S is also described in [7], and is given by

$$S = \begin{bmatrix} 1 & & & & \\ 1 & 1 & & & \\ 1 & 2 & 1 & & \\ \vdots & \vdots & \vdots & \ddots & \\ \binom{r+1}{0} & \binom{r+1}{1} & \binom{r+1}{2} & \cdots & 1 \end{bmatrix}$$

This matrix is a special case of a more general form

$$M(\lambda) = \left(\binom{i - 1}{j - 1}\lambda^{i-j}\right)_{ij}.$$

In fact, $M(\lambda)$ plays the role of S when the characteristic polynomial is $(t - \lambda)^{r+2}$, and the specific instance of S above is $M(1)$. There are several interesting properties of $M(\lambda)$ described in [6]. Of special interest here is that

$$M(\lambda)^{-1} = M(-\lambda),$$

and in particular,

$$S^{-1} = M(-1).$$

This shows that S^{-1} has the same form as S, but with a minus sign introduced before each entry of the odd numbered subdiagonals. Note that the square matrix which appears in (2) has exactly this form, but with one less row and column. Put another way, the matrix in (2) is the $(r + 1)$-dimensional principal submatrix of S^{-1}. For future reference, we shall denote this matrix by S*.

Combining the results presented so far, we have

$$(10) \qquad s_n^r = [1 \ \ 0 \ \ 0 \ \cdots \ 0]SJ^nS^{-1}[s_0^r \ \ s_1^r \ \ s_2^r \ \cdots \ s_{r+1}^r]^{\mathrm{T}}.$$

This equation can be simplified by observing that premultiplication of a square matrix by the row $[1 \ \ 0 \ \ 0 \ \cdots \ 0]$ results in just the first row of the matrix. The first row of S is again $[1 \ \ 0 \ \ 0 \ \cdots \ 0]$ so that the product of the first three factors on the right side of (10) is simply the first row of J^n, or

$$\left[\binom{n}{0} \ \ \binom{n}{1} \ \ \binom{n}{2} \ \cdots \ \binom{n}{r+1}\right].$$

Therefore, we may write

$$(11) \qquad s_n^r = \left[\binom{n}{0} \ \ \binom{n}{1} \ \ \binom{n}{2} \ \cdots \ \binom{n}{r+1}\right]S^{-1}[s_0^r \ \ s_1^r \ \ s_2^r \ \cdots \ s_{r+1}^r]^{\mathrm{T}}$$

This is similar to (2), and is interesting in its own right.

Next, to replace the initial terms of the sequence $\{s_n^r\}$ with initial terms of $\{n^r\}$, we observe that

$$(12) \qquad \begin{bmatrix} s_0^r \\ s_1^r \\ s_2^r \\ \vdots \\ s_{r+1}^r \end{bmatrix} = \begin{bmatrix} 1 & & & & \\ 1 & 1 & & & \\ 1 & 1 & 1 & & \\ \vdots & \vdots & \vdots & \ddots & \\ 1 & 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} 0^r \\ 1^r \\ 2^r \\ \vdots \\ (r+1)^r \end{bmatrix}.$$

When the right side of (12) is substituted in (11), the product $S^{-1}T$ appears, where T is the triangular matrix in (12). A straightforward computation reveals that $S^{-1}T$ may be expressed as the partitioned matrix

$$\left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & S^* \end{array}\right].$$

Thus, the combination of (11) and (12) results in the partitioned matrix equation

$$(13) \qquad s_n^r = \left[\binom{n}{0} \middle| \binom{n}{1} \ \ \binom{n}{2} \ \cdots \ \binom{n}{r+1}\right] \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & S^* \end{array}\right] [0 \,|\, 1^r \ \ 2^r \ \cdots \ (r+1)^r]^{\mathrm{T}}.$$

Carrying out the partitioned multiplication completes the derivation of (2).

It is instructive to use (2) to derive a formula for s_n^r in a particular case. For example, with $r = 2$, we have

$$s_n^2 = \left[\binom{n}{1} \; \binom{n}{2} \; \binom{n}{3}\right] \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \\ 9 \end{bmatrix} = \binom{n}{1} + 3\binom{n}{2} + 2\binom{n}{3}.$$

This gives s_n^2 in terms of binomial coefficients, and simplification produces the well-known equation

$$s_n^2 = \frac{n(n+1)(2n+1)}{6}.$$

The derivation of (2) generalizes immediately. Let p be a polynomial of degree r, and define $s_n = \sum_{k=0}^{n} p(k)$. All of the analysis through equation (13) remains valid when k^r is replaced by $p(k)$. This leads to the following analog of equation (13).

$$(14) \qquad s_n = \left[\binom{n}{0} \middle| \binom{n}{1} \; \binom{n}{2} \; \cdots \; \binom{n}{r+1}\right] \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & S^* \end{array}\right] [p(0) \,|\, p(1) \;\; p(2) \;\; \cdots \;\; p(r+1)]^{T}$$

Carrying out the partitioned product now yields the identity

$$(15) \qquad s_n - p(0) = \left[\binom{n}{1} \; \binom{n}{2} \; \binom{n}{3} \; \cdots \; \binom{n}{r+1}\right] \begin{bmatrix} 1 & & & & \\ -1 & 1 & & & \\ 1 & -2 & 1 & & \\ \vdots & \vdots & \vdots & & \\ \pm 1 & \mp\binom{r}{1} & \pm\binom{r}{2} & \cdots & 1 \end{bmatrix} \begin{bmatrix} p(1) \\ p(2) \\ p(3) \\ \vdots \\ p(r+1) \end{bmatrix}.$$

This equation may be used for adding up the first n terms of the sequence $\{p(k)\}$ starting from $k = 1$ instead of $k = 0$.

An interesting class of examples involves summing the r^{th} powers of the first n integers equivalent to b modulo a. In these cases, the polynomial has the form $p(k) = (ak + b)^r$. With $a = 4$ and $b = -3$, for example, the left-hand side of (15) is the sum of the first n terms of the progression $1^r, 5^r, 9^r, \ldots$. For an even more specific example, let $r = 2$. Then (15) reduces to

$$1^2 + 5^2 + 9^2 + \cdots + (4n-3)^2 = \left[\binom{n}{1} \; \binom{n}{2} \; \binom{n}{3}\right] \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1^2 \\ 5^2 \\ 9^2 \end{bmatrix}$$

$$= \binom{n}{1} + 24\binom{n}{2} + 32\binom{n}{3}.$$

A review of the derivation of (2) and (15) reveals a natural division into two parts. In the first, culminating in equation (11), the analysis has general validity. Any difference equation for which the Jordan canonical form can be calculated can be subjected to a similar analysis, resulting in an analogous identity. The second part depends on the fact that the characteristic polynomial for the difference operator is a power of $t - 1$. Therefore, the final result (15) should not be expected to generalize in any obvious fashion to a larger class of difference equations. In the final section, another example is considered. As expected, a result analogous to (11) is obtained, but no analog for (2) appears.

4. Geometrically Weighted Power Sums

In [3], recursive procedures are presented for expressing formulas for the geometrically weighted power sum

$$s_n^r(x) = \sum_{k=0}^{n} k^r x^k.$$

This is a generalization of s_n^r in the sense that $s_n^r(1) = s_n^r$. The sequence $\{s_n^r(x)\}$ (indexed by n) can be analyzed by matrix difference equation methods. As a first step, we have the following simple generalizations of earlier material.

Definition 3: D_x is the operator $L - x$, where x acts as a scalar multiplier. $N_k(x)$ is the null space of D_x^k. Note that $N_1(x) \subseteq N_2(x) \subseteq N_3(x) \ldots$.

Theorem 2: For any k, $N_k(x)$ consists of the sequences $\{a_n\}$ defined as the termwise product of the exponential sequence x^n with a polynomial in n of degree less than k.

We omit a proof for this theorem; one can be obtained by modifying the proof of the earlier theorem in an obvious way. The main significance for the present discussion is as follows. Since

$$Ds_n^r(x) = (n+1)^r x^{n+1},$$

it must be annihilated by D_x^{r+1}. Therefore, $s_n^r(x)$ is a solution to the difference equation

$$(L - x)^{r+1}(L - 1)a_n = 0.$$

The characteristic polynomial for this equation is $(t - x)^{r+1}(t - 1)$. We represent the reduction of its companion matrix to Jordan form in the usual way as $C = SJS^{-1}$. Once again, the analysis of [7] is directly applicable. It tells us that J has one Jordan block of dimension $r + 1$ for the root x, and a 1×1 block for the simple root 1. The matrix S is closely related to $M(x)$ defined above. In fact, the first $r + 1$ columns of S are identical to the corresponding columns of $M(x)$, but the final column consists of all 1's. [Indeed, this final column is really the first column of $M(1)$. In general, the matrix S is a combination of $M(x)$'s for the various roots of the characteristic polynomial, with the number of columns for each x given by its multiplicity as a root.] With these definitions for J and S, and with $s_n^r(x)$ in place of s_n^r, we may calculate $s_n^r(x)$ using (11).

Unfortunately, there is a bit more work required to determine an explicit representation for the inverse of S in this example. For simplicity, shorten $M(x)$ to M, and define E to be the difference S - M. Thus, E is given by

$$\begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 0 & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 0 \end{bmatrix} [0 \quad 0 \quad \cdots \quad 0 \quad 1].$$

Let the column and row matrices in this factorization be called E_c and E_r, respectively. Now we claim that

$$S^{-1} = (I - S^{-1}E)M^{-1}.$$

This can be verified by premultiplying by S. The product $S^{-1}E$ can be computed using the factorization of E as soon as $S^{-1}E_c$ is determined. Thus, the problem of inverting S is reduced to finding the inverse image of a single vector. This is not surprising: since S was obtained by making a rank 1 modification to M, it is reasonable to expect a corresponding rank 1 modification to link the inverse matrices.

Proceeding with this approach, $S^{-1}E_c$ is computed by solving the equation $Sv = E_c$ for v. Again using $S = M + E_cE_r$, write the equation as

$$Mv + E_cE_rv = E_c.$$

Since E_rv is a scalar, namely v, the last entry of v, the equation may now be rearranged as

$$Mv = (1 - v)E_c.$$

This leads to

$$v = (1 - v)M^{-1}E_c,$$

and by equating the final entries of the vectors on either side, to an equation for v. Once v is found, v simply requires the computation shown at the right side of the previous equation. Carrying out these steps produces

$$v = \frac{1}{(1 - x)^{r+1}}\begin{bmatrix} 1 \\ 1 - x \\ (1 - x)^2 \\ \vdots \\ (1 - x)^r \\ (1 - x)^{r+1} - 1 \end{bmatrix}$$

With this result, it is now possible to express S^{-1} as $M^{-1} - vE_rM^{-1}$. Let $w = E_rM^{-1}$, which is simply the last row of M^{-1}. This gives $S^{-1} = M^{-1} - vw$.

At this point, the factors appearing at right in (11) cannot be simplified much further. As before, the first two factors yield the first row of S. However, this row has a 1 in the last position as well as the first, so multiplying by J^n results in the sum of first and last rows of that matrix. Meanwhile, J is a block diagonal matrix. The first block is $(r + 1)$-dimensional and of the form $xI + N$. Its powers are computed just as before, exploiting the properties of N. Specifically, the first row of the n^{th} power is

$$\left[\binom{n}{0}x^n \quad \binom{n}{1}x^{n-1} \quad \cdots \quad \binom{n}{r}x^{n-r}\right]$$

and contributes all but the last entry of the first row of J^n. The second block is just the scalar 1 at the end of the diagonal. It contributes the only nonzero entry in the last row of J^n. When the first and last row are added, the result is

$$\left[\binom{n}{0}x^n \quad \binom{n}{1}x^{n-1} \quad \cdots \quad \binom{n}{r}x^{n-r} \quad 1\right].$$

When all of the foregoing calculations and reductions are combined into a single equation, the result is

$$(16) \qquad s_n^r(x) = \left[\binom{n}{0}x^n \quad \binom{n}{1}x^{n-1} \quad \cdots \quad \binom{n}{r}x^{n-r} \quad 1\right]S^{-1}\begin{bmatrix} s_1^r(x) \\ s_1^r(x) \\ s_2^r(x) \\ \vdots \\ s_{r+1}^r(x) \end{bmatrix}$$

where

$$
S^{-1} = \begin{bmatrix} 1 & & & & \\ -x & 1 & & & \\ x^2 & -2x & 1 & & \\ \vdots & \vdots & \vdots & \ddots & \\ (-x)^{r+1} & \binom{r+1}{1}(-x)^r & \binom{r+1}{2}(-x)^{r-1} & \dots & 1 \end{bmatrix}
$$

$$
- \frac{1}{(1-x)^{r+1}} \begin{bmatrix} 1 \\ 1-x \\ \vdots \\ (1-x)^r \\ (1-x)^{r+1} - 1 \end{bmatrix} \begin{bmatrix} (-x)^{r+1} & \binom{r+1}{1}(-x)^r & \binom{r+1}{2}(-x)^{r-1} & \dots & 1 \end{bmatrix}
$$

This formulation is not as compact as (2) but is sufficiently orderly to permit convenient calculation for specific values of r and x. The following formulas were obtained by writing a short computer program to define and calculate the product of the last two matrices on the right side of (16), then running it with x set to 2 and r set to 1, 2, 3, 4, and 5.

$$s_n^1(2) = 2^n\left[-2 + 2\binom{n}{1}\right] + 2$$

$$s_n^2(2) = 2^n\left[6 - 2\binom{n}{1} + 4\binom{n}{2}\right] - 6$$

$$s_n^3(2) = 2^n\left[-26 + 14\binom{n}{1} + 24\binom{n}{3}\right] + 26$$

$$s_n^4(2) = 2^n\left[150 - 74\binom{n}{1} + 52\binom{n}{2} + 24\binom{n}{3} + 48\binom{n}{4}\right] - 150$$

$$s_n^5(2) = 2^n\left[-1082 + 542\binom{n}{1} - 240\binom{n}{2} + 300\binom{n}{3} + 250\binom{n}{4} + 240\binom{n}{5}\right] + 1082$$

These equations are similar to the ones derived by Gauthier ([3], eqs. 31), but express $s_n^r(2)$ in terms of binomial coefficients instead of as polynomials in n.

It is also feasible to use (16) symbolically for small values of r. As an example, we carry through the matrix multiplication for $r = 2$.

The algebra will be simplified if the factors of $(1 - x)$ appearing in the denominator of entries of S^{-1} are transferred to the corresponding entries of the first matrix factor. In pursuit of this goal, rewrite (16) in the form

$$s_n^2(x) = RS^{-1}C,$$

where R and C are the row and column vectors, respectively, appearing in (16). Next, define the diagonal matrix D with entries

$$(1-x)^{-3},\ (1-x)^{-2},\ (1-x)^{-1},\ \text{and}\ (1-x)^{-3}.$$

Then we may write (17) as

$$s_n^2(x) = (RD)(D^{-1}S^{-1})C.$$

Focusing separately on each factor in (18), observe that

$$
\begin{aligned}
RD &= \left[\binom{n}{0}\frac{x^n}{(1-x)^3} \quad \binom{n}{1}\frac{x^{n-1}}{(1-x)^2} \quad \binom{n}{2}\frac{x^{n-2}}{1-x} \quad \frac{1}{(1-x)^3}\right] \\
C^T &= \left[\quad 0 \quad x \quad 4x^2 + x \quad 9x^3 + 4x^2 + x\right]
\end{aligned}
$$

and

$$\mathbf{D}^{-1}\mathbf{S}^{-1} = \begin{bmatrix} (1-x)^3 & & & \\ -x(1-x)^2 & (1-x)^2 & & \\ x^2(1-x) & -2x(1-x) & 1-x & \\ -x^3(1-x)^3 & 3x^2(1-x)^3 & -3x(1-x)^3 & (1-x)^3 \end{bmatrix}$$

$$- \begin{bmatrix} 1 \\ 1 \\ 1 \\ (1-x)^3 - 1 \end{bmatrix} [-x^3 \quad 3x^2 \quad -3x \quad 1].$$

Expressing the right side of this equation as a single matrix produces

$$\mathbf{D}^{-1}\mathbf{S}^{-1} = \begin{bmatrix} * & -3x^2 & 3x & -1 \\ * & -2x^2 - 2x + 1 & 3x & -1 \\ * & -x^2 - 2x & 2x + 1 & -1 \\ * & 3x^2 & -3x^2 & 1 \end{bmatrix}.$$

The entries in the first column have not been explicitly presented because they have no effect on the final formula for $s^2(x)$; these entries are each multiplied by the zero in the first position of $\mathbf{C}$. Indeed, multiplying this last expression by $\mathbf{C}$ now yields

$$\mathbf{D}^{-1}\mathbf{S}^{-1}\mathbf{C} = \begin{bmatrix} -x(x+1) \\ x^2(x-3) \\ -2x^3 \\ x(x+1) \end{bmatrix}.$$

Finally, after multiplying by $\mathbf{RD}$, the following formula is obtained:

$$s_n^2(x) = x^{n+1}\left[\frac{-(x+1)}{(1-x)^3} + \binom{n}{1}\frac{x-3}{(1-x)^2} - \binom{n}{2}\frac{2}{1-x}\right] + \frac{x^2+x}{(1-x)^3}$$

As before, this result is consistent with the analysis presented in [3].

5. Summary

In this paper, matrix methods have been used to derive closed form expressions for the solutions of difference equations. The general tool of analysis involves expressing a scalar difference equation of order k as a first-order vector equation, then using the Jordan canonical form to express powers of the system matrix, thus describing the solution to the equation. Two specific examples of the method have been presented, differing from previous work in that neither example features a diagonalizable system matrix. In the first example, an esthetically appealing formula for the sum $\sum_{k=0}^{n} k^r$ was derived. In the second example, the more general sum $\sum_{k=0}^{n} k^r x^k$ was analyzed. In each case, the results have been derived previously using other methods. However, the main point of the article has been to show that the methods of matrix algebra can be a powerful tool, and provide a distinct heuristic insight, for the study of difference equations.

References

1. Tom M. Apostol. *Calculus*, Vol. II, 2nd ed. Waltham, Mass.: Blaisdell, 1969.
2. M. C. Er. "Matrices of Fibonacci Numbers," *Fibonacci Quarterly 22.2* (1984): 134-139.

3. N. Gauthier. "Derivation of a Formula for $\sum r^k x^r$." *Fibonacci Quarterly* *27.5* (1989):402-408.
4. H. W. Gould. *Sums of Powers of Integers*, volume Part I of *Number Theory Class Notes*. West Virginia University, 1974-1975.
5. Dan Kalman. "Generalized Fibonacci Numbers by Matrix Methods." *Fibonacci Quarterly* *20.1* (1982):73-76.
6. Dan Kalman. "Polynomial Translation Groups." *Math. Mag.* *56.1* (1983):23-25.
7. Dan Kalman. "The Generalized Vandermonde Matrix." *Math. Mag.* *57.1* (1984): 15-21.
8. Kenneth A. Miller. *Linear Difference Equations*. New York: Benjamin, 1968.
9. Evar D. Nering. *Linear Algebra and Matrix Theory*, 2nd ed. New York: Wiley, 1970.
10. Jeffrey Nunemacher & Robert M. Young. "On the Sum of Consecutive k^{th} Powers." *Math. Mag.* *60.4* (1987):237-238.
11. J. L. Paul. "On the Sum of the k^{th} Powers of the First n Integers." *Amer. Monthly* *78.3* (1971):271-272.
12. Henry J. Schultz. "The Sum of the k^{th} Powers of the First n Integers." *Amer. Math. Monthly* *87.6* (1980):478-481.
13. Barbara Turner. "Sums of Powers of Integers via the Binomial Theorem." *Math. Mag.* *53.2* (1980):92-96.