The next number in the sequence

JOHN D. MAHONY

Readers familiar with numbers and their patterns will most likely know of the game-playing exercise where one is invited to spot the next number in a sequence, only to find that after giving an answer there can be a multiplicity of possibilities, each of which is valid according to its own rule of sequence construction. This is an instance of partial sequence matching and the purpose of this note is to show how a simple method of constructing a partial sequence match can be achieved. It requires only a knowledge of polynomials and matrix algebra, which are topics within the grasp of an undergraduate, and will be instanced against the backdrop of some well-known sequences.

The concept of a sequence is well known to mathematicians and it is simply a collection of terms, which is ordered according to some rule or procedure. In the case of an integer sequence, one can be more precise by stating that an integer sequence is essentially a mapping from $\mathbb{N}$ (the set of natural numbers, or, positive integers) to $\mathbb{Z}$ (the set of integers). Much about them can be gleaned from the internet by googling the appropriate words. The mapping can be described by a relationship that defines uniquely each term in the sequence. For example, the value of the term in the sequence can be denoted by I_n $(n = 1, 2, 3, \dots)$ where I_n is a function of the parameter n. Alternatively, the terms of the sequence can be defined inductively, recursively, or in a definite mathematical context, such as discussed below in the first case.

With the above terminology, a finite subset of consecutive terms in a sequence will be denoted by $\{I_n, I_{n+1}, \dots, I_{n+m}\}$, where n is some natural number. If $n = 1$, which will be the case here later, the subset is simply the first $(m + 1)$ terms of the sequence. In general, to achieve a polynomial match to this subset it will suffice first to assume a polynomial, $F(x)$, in the form

$$F(x) = I_n + \sum_{i=1}^{m} A_i (x - n)^i \qquad (1)$$

where x assumes in turn each value in the range $n, n + 1, \dots, n + m$. When $x = n$ (the starting value for the subset), the polynomial value is simply $F(n) = I_n$. If q denotes some other x-value in this range, an expression involving the coefficients A_i $(i = 1, 2, 3, \dots, m)$ can be determined by forcing the match $F(q) = I_q$. For all such possible q-values $(q = n + 1, \dots, n + m)$, this results in the following matrix system of m simultaneous equations to determine the A_i:

https://doi.org/10.1017/mag.2024.108 Published online by Cambridge University Press

$$\begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 2 & 2^2 & 2^3 & \dots & 2^m \\ 3 & 3^2 & 3^3 & \dots & 3^m \\ 4 & 4^2 & 4^3 & \dots & 4^m \\ & & \vdots & & \\ m & m^2 & m^3 & \dots & m^m \end{pmatrix} \begin{pmatrix} A_1 \\ A_2 \\ A_3 \\ A_4 \\ \vdots \\ A_m \end{pmatrix} = \begin{pmatrix} I_{n+1} - I_n \\ I_{n+2} - I_n \\ I_{n+3} - I_n \\ I_{n+4} - I_n \\ \vdots \\ I_{n+m} - I_n \end{pmatrix} \qquad (2)$$

This system of equations can be solved directly using the numerical method of pivotal condensation [1] or Gaussian elimination [2]. The procedures are straightforward and are well suited to classroom practice when the coefficient matrix in (2) is not large. Similarly, a matrix inversion approach using Jordan's method can be employed [2] to determine the inverse of the coefficient matrix. Fortunately, a matrix inversion method is built into the software on an excel spreadsheet and it is convenient to use this also for intermediate-sized coefficient matrices, provided of course that the methods are fully understood. It is thus possible to determine the A_i numerically, and hence the value of the polynomial given in (1) can be computed for various values of the input parameter x. At those values where a match has been forced, results from the polynomial evaluation will perforce reproduce the sequence subset values.

Having proposed a way of reproducing a partial subset of consecutive terms of a sequence, we will focus now on the practice as it relates to the first few terms of some well-known sequences.

1. *As a paradigm we consider the number sequence* $\{1, 2, 4, 8, 16, \dots\}$: When this sequence is put before young students without further ado and with an entreaty to guess the next number, the obvious candidate is usually assumed to be 32, because each number in the sequence appears to be double that of the preceding one. This is certainly the case if the sequence is defined to be $\{2^{n-1}; n = 1, 2, 3, 4, 5, \dots\}$. However, the Canadian Mathematician. Leo Moser, gave a method of sequence construction which produced an alternative sequence in which the sixth number was 31. His example concerned the maximum number of regions into which chords joining n points on a circle's circumference partitioned its enclosed area. An exposition can be found in a 'You-Tube' presentation [3]. Briefly, arguments there were based on combinatorics and discrete mathematics to develop a binomial expression yielding the sequence $\{1, 2, 4, 8, 16, 31, 57, 99, \dots\}$, which starts the same as the example but ends up differently after the fifth term of the sequence. This, and more, can be appreciated by googling the phrase "Moser intersecting chords" to see a number of internet leads on the subject, together with a variety of similar 'You-Tube' presentations. Generally, the argument exploits the well-known Euler characteristic formula that relates the number of vertices (V), edges (E) and faces (F) on straight-edged shapes, viz. $V - E + F = 2$, in tandem

with appreciations concerning number structures in a Pascal triangle. Material that is relevant to such matters can be found in [4, 5, 6, 7]. Whilst this argument leads to the development of a fourth order polynomial in the guise of a binomial expression, it is possible to secure the same result in a different way by adopting the above-mentioned protocol of a polynomial fit to the data. This approach will be considered here for the class of sequence we use as a paradigm sequence, i.e.

$$\{p^{n-1};\ n = 1, 2, 3, 4, 5, \dots\}, \tag{3}$$

where p is some positive integer. The case of $p = 2$ is the above paradigm sequence. To illustrate the procedure, the details of a polynomial fit to the first five terms of the sequence in (3) (with $n = 1$) will be discussed. In this case, $m = 4$ and so $I_1 = 1$, $I_2 = 2$, $I_3 = 4$, $I_4 = 8$ and $I_5 = 16$. The corresponding matrix equation system given by (2) is then

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & 2^2 & 2^3 & 2^4 \\ 3 & 3^2 & 3^3 & 3^4 \\ 4 & 4^2 & 4^3 & 4^4 \end{pmatrix} \begin{pmatrix} A_1 \\ A_2 \\ A_3 \\ A_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 7 \\ 15 \end{pmatrix}.$$

A solution for the A_i, obtained in the above described fashion, yields the solutions $A_1 = \frac{7}{12}$, $A_2 = \frac{11}{24}$, $A_3 = -\frac{1}{12}$, $A_4 = \frac{1}{24}$. The corresponding polynomial expression from (1), with $(x - 1)$ replaced by τ, can then be written

$$F(x) = 1 + \frac{7}{12}\tau + \frac{11}{24}\tau^2 - \frac{1}{12}\tau^3 + \frac{1}{24}\tau^4. \tag{4}$$

It is now a simple matter to verify that the x-values of 1, 2, 3, 4, 5, 6, 7, 8, …, respectively, produce from this polynomial the sequence of numbers $\{1, 2, 4, 8, 16, 31, 57, 99, \dots\}$. The first five terms of this sequence are, as expected, the same as the first five numbers in the paradigm sequence. The above polynomial can be compared with the commonly accepted binomial expression given in [1], i.e. $\left(\binom{x}{4} + \binom{x}{2} + 1\right)$. It is left as an exercise for the reader to appreciate that the two results are identical. At the first point where there is a mismatch in values between respective terms in the two sequences (the paradigm one and the polynomial one), the difference is unity $(= 32 - 31)$. It is of interest to note also that the polynomial expression to which this binomial reduces is the same as that to which the binomial expression $\left\{\binom{n-1}{0} + \binom{n-1}{1} + \binom{n-1}{2} + \binom{n-1}{3} + \binom{n-1}{4}\right\}$ also reduces, when n is replaced by x. The latter represents the number of regions in 4-space formed by hyperplanes.

The above-described numerical procedure can be developed to incorporate more terms in the polynomial match so as to extend the level of

https://doi.org/10.1017/mag.2024.108 Published online by Cambridge University Press

agreement between the terms in the polynomial sequence and terms in the given candidate sequence. For example, if the Number of Consecutive Terms (NCT values) over which a sequence match is required (between the candidate sequence and the polynomial one) is $m + 1$, then a polynomial of order m will suffice. Computations were carried out to determine the coefficients A_i for a variety of different NCT values pertaining to the current candidate sequence. They are shown in Table 1.

NCT	A_1 / G	A_2 / G	A_3 / G	A_4 / G	A_5 / G	A_6 / G	A_7 / G	G
2	1	-	-	-	-	-	-	1
3	1	1	-	-	-	-	-	$\frac{1}{2}$
4	5	0	1	-	-	-	-	$\frac{1}{6}$
5	14	11	−2	1	-	-	-	$\frac{1}{24}$
6	94	5	25	−5	1	-	-	$\frac{1}{120}$
7	444	304	−75	55	−9	1	-	$\frac{1}{720}$
8	3828	364	1099	−350	112	−14	1	$\frac{1}{5040}$

TABLE 1: Coefficient values required to achieve a partial match of sequences at the given number of terms when $p = 2$.

The results in the table have been presented in normalised form, and there seems to be a predictable pattern for the normalising factor G. Results for the coefficients in (4) above are captured when the NCT value is 5. A further example to consider on the use of the Table is when the value of NCT is 7 (the second to last row in the Table). The normalising factor in this instance is $\frac{1}{720}$ and the value of the coefficients required in the corresponding polynomial of order 6 can be selected from the relevant row in the Table for use in (1). With $(x - 1)$ replaced by τ, the associated polynomial result is

$$F(\tau) = 1 + \frac{1}{720}\left(444\tau + 304\tau^2 - 75\tau^3 + 55\tau^4 - 9\tau^5 + \tau^6\right).$$

It is a straightforward matter to verify that, for x-integer values given respectively by $1, 2, 3, 4, 5, 6, 7, 8, 9, \ldots,$ this polynomial expression yields the corresponding sequence values $\{1, 2, 4, 8, 16, 32, 127, 247, \ldots\}$. The first seven terms here match those of the paradigm case. Again, there is a delta value of unity $(= 128 - 127)$ at the first point where there is a mismatch between terms in the paradigm and polynomial sequences.

Other sequences in the paradigm class: By way of illustration, the case of $p = 3$ in (3) will be considered when $m = 6$ in (1); that is to say, there will be a forced fit between the first 7 $(= m + 1)$ terms of the specified sequence and that of output from the polynomial. The terms of this

https://doi.org/10.1017/mag.2024.108 Published online by Cambridge University Press

paradigm class sequence are $\{1, 3, 9, 27, 81, 243, 729, 2187, 6561, \dots\}$. Again using the established spreadsheet facilities, the coefficients A_i for $i = 1, 2, 3, 4, 5, 6$ were found to be as shown in Table 2.

NCT	A_1/G	A_2/G	A_3/G	A_4/G	A_5/G	A_6/G	G
7	−252	736	−600	250	−48	4	$\frac{1}{45}$

TABLE 2: Values of coefficients in the polynomial when $p = 3$ and $m = 6$

With $(x - 1)$ again replaced by τ in the polynomial expression of (1), the corresponding polynomial in this case can be written

$$F(x) = 1 + \frac{1}{45}\left(-252\tau + 736\tau^2 - 600\tau^3 + 250\tau^4 - 48\tau^5 + 4\tau^6\right).$$

Sequence values obtained using this polynomial for $x = 1, 2, 3, \dots, 9, \dots$, were found to be $\{1, 3, 9, 27, 81, 243, 729, 2059, 5281, \dots\}$. As expected, the first seven values of this sequence are identical to the first seven values above. At the first instance of a mismatch in respective sequence terms, the delta value is 128 $(= 2187 - 2059)$. In the earlier case when $p = 2$, such a difference was found to be unity. After examining a number of cases, it seems that such differences at the first point of mismatch for the class of sequences under consideration (see (3)) can be described by the formula $(p - 1)^{m+1}$.

It is left as an exercise for the reader to explore agreements between any number of sequence values determined using the polynomial matching equation of (1) for other values of p, but the required numerical effort increases significantly as such values and numbers increase in size.

2. *Another sequence of interest*: Also of interest is a match to the floor function sequence

$$\left\{INT\left(2^{n-2}\right);\ n = 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, \dots\right\}$$

which produces respectively the sequence values

$$0, 1, 2, 4, 8, 16, 32, 64, 128, 512, \dots .$$

With the exception of an additional first term of zero, the other terms here are the same as those in the first paradigm sequence above. The 10th term of this sequence is 256. A polynomial match over various values can be achieved with the coefficients shown in Table 3.

https://doi.org/10.1017/mag.2024.108 Published online by Cambridge University Press

NCT	A_1/G	A_2/G	A_3/G	A_4/G	A_5/G	A_6/G	A_7/G	G
3	1	-	-	-	-	-	-	1
4	8	-3	1	-	-	-	-	$\frac{1}{6}$
5	8	-3	1	0	-	-	-	$\frac{1}{6}$
6	184	-110	55	-10	1	-	-	$\frac{1}{120}$
7	184	-110	55	-10	1	0	-	$\frac{1}{120}$
8	8448	-6384	3934	-1155	217	-21	1	$\frac{1}{5040}$
9	8448	-6384	3934	-1155	217	-21	1	$\frac{1}{5040}$

TABLE 3: Polynomial coefficient values required to match floor function sequence values at various NCT values

It is left as an exercise for the reader to determine explicitly the associated polynomials, and it is apparent that matches at some NCT values suffice also for a match at the next one.

3. *A further type of integer sequence*: The above described method of fitting a polynomial to selected consecutive terms of a given sequence can be applied to yet other known sequences. For example, the Sloane sequence $\{1, 1, 2, 4, 8, 16, 30, 60, 96, 160, 270, \dots\}$ [8] is of interest. This sequence is formed by determining the number of whole number divisors of $(x - 1)!$ when x assumes, respectively, the values 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11. Again, it is possible in the usual way to determine coefficients in a matching polynomial that depend on the NCTs in the Sloane sequence that are to be matched. Typical results are shown in Table 4.

NCT	A_1	A_2	A_3	A_4	A_5	A_6	A_7
3	$-\frac{1}{2}$	$\frac{1}{2}$	-	-	-	-	-
4	$-\frac{1}{2}$	$\frac{1}{2}$	0	-	-	-	-
5	$-\frac{3}{4}$	$\frac{23}{24}$	$-\frac{1}{4}$	$\frac{1}{24}$	-	-	-
6	$-\frac{3}{4}$	$\frac{23}{24}$	$-\frac{1}{4}$	$\frac{1}{24}$	0	-	-
7	$-\frac{7}{12}$	$\frac{26}{45}$	$\frac{1}{16}$	$-\frac{11}{144}$	$\frac{1}{48}$	$-\frac{1}{720}$	-
8	$\frac{71}{84}$	$-\frac{263}{90}$	$\frac{473}{144}$	$-\frac{221}{144}$	$\frac{53}{144}$	$-\frac{31}{720}$	$\frac{1}{504}$

TABLE 4: Polynomial coefficient values required for matching various terms in the Sloane sequence

Once more, it is interesting to note that matches at some NCT values suffice also for a match at the next NCT value. Again, details of the matching polynomial are left for the reader to explore.

https://doi.org/10.1017/mag.2024.108 Published online by Cambridge University Press

Discussion and conclusion: The concept about the meaning of a pattern in sequences is ubiquitous. People see trends, patterns and themes (and their variations) in all facets of human endeavour, be it in the arts (especially music), the sciences, mathematics or other disciplines. The search for understanding and appreciating them can move the state of knowledge in any such discipline forward. Particularly, spotting pattern behaviour in sequences is an endeavour that should be encouraged in a mathematician, whether of the pure or applied variety. When mathematicians view patterns they might come with preconceptions, or not. Whichever, established ways of viewing things might, as a result, be skewed for the good. The pattern one sees in a sequence when a prescription is not on offer is an invitation to extend a repertoire, it is a force majeure that serves to stimulate the enquiring mind much like musical variations that form in the mind to complete the ending to a piece that is different to that envisaged by the composer. Whenever a next-number candidate in a sequence has been spotted, it should be appreciated by now that there will be other means that produce a different candidate. Without knowing how a given integer sequence is ordered, a next term could be almost anything. The above polynomial matching procedure might, within reason, be employed to determine it as long as the 'anything' is incorporated in the process. For instance, in the above example due to Moser it was demonstrated that the sixth term in the sequence was 31, not 32 as some might have expected. However, using the above described procedure, it is possible to capture the first five terms of Moser's sequence and for the 6th term to be, say, 33. The 5th order polynomial that will achieve as much is

$$\left(-1 + \frac{229}{60}x - \frac{67}{24}x^2 + \frac{7}{6}x^3 - \frac{5}{24}x^4 + \frac{1}{60}x^5\right).$$

It is left as an exercise for the reader to verify this.

The reader might appreciate also that the unexpected can occur in pattern-spotting even when the terms in a sequence are well-defined. One need look no further than the behaviour of coefficients in the 105th cyclotomic polynomial. In lesser polynomials of this type, the coefficients are generally ± 1 and/or 0, but in the 105th polynomial the number 2 appears unexpectedly as one of the coefficients (see [9]).

Finally, the process of matching output from a polynomial to some consecutive terms in a target integer sequence has resulted in rational coefficients in the matching polynomial, but it remains to be seen if subsequent output terms from the polynomial sequence (terms where a match has not been forced) will always be integers.

A standard example of spotting the next number in an integer sequence has been re-examined in the light of matching a polynomial to various paradigm cases, and the argument has been extended to consider other types of integer sequences. There are numerous other examples that can be considered in the same light. The polynomial matching process has been applied to match only the first few terms of a candidate sequence, and it is well suited to such cases. However, if the number of terms to be matched is increased, the ensuing numbers in the associated matrix in (2) become large very quickly and the subsequent matrix inversion process involves numbers that likewise become small with increasing matrix size. Similarly, the power terms in the matching polynomial become large quickly as their arguments increase with the number of sequence terms to be matched.

Acknowledgements

The author is pleased to acknowledge helpful and useful comments from a referee.

References

1. National Physical Laboratory, *Notes on applied science no. 16, Modern computing methods*, Her Majesty's Stationary Office, 1961.
2. Carl-Erik Froberg, *Introduction to numerical analysis,* (2nd edn.), Addison-Wesley, (1970).
3. Circle division solution, *YouTube*, accessed March 2024 at https://www.youtube.com/watch?v=K8P8uFahAgc
4. Leo Moser and W. Bruce Ross, On the Danger of Induction, *Mathematics Magazine*, **23** (2) (November – December, 1949), pp. 109-114.
5. R. K. Guy, The strong law of small numbers, *Amer. Math. Monthly*, **95** (8) (1988) pp. 697-712.
6. R. K. Guy, The second strong law of small numbers, *Math. Magazine*, **63** (1990), pp. 3-20.
7. D. MacHale, *Number and letter sequences*, Logic Press (2019).
8. Glen Burch, Number of divisors of n!, OEIS, accessed March 2024 at *https: // oeis.org*
9. Cyclotomic polynomials, *Wikipedia*, accessed March 2024 at *https: // en.wikipedia.org/wiki/cyclotomic_polynomials*

10.1017/mag.2024.108
Published by Cambridge University Press
on behalf of The Mathematical Association

JOHN D. MAHONY
5 Bluewater View, Mt. Pleasant, Christchurch 8081, New Zealand
e-mail: *johndmahony@gmail.com*