

1

Difference Tables and Polynomial Fits

Introduction

Many curricula ask students to look for patterns in data. Often, students are given a table like this:

Input	Output
0	3
1	8
2	13
3	18
4	23
5	28
6	33

And they are asked to find a “formula” that produces the table.

Many beginning students don’t look for a “closed form” solution at first. They don’t ask “what can I do to 5 to get 28 so that when I do the same things to 6, I get 33?” Instead, they often subtract successive outputs and notice that the outputs “go up by 5 each time.” It turns out that, if the differences go up by the same amount, there’s a linear function (in this case, $f(x) = 5x + 3$) that fits the table. We’ll see why in the next section.

Sometimes, the differences between successive outputs are *not* constant, but the successive differences of the differences *are* constant. When the “second differences” are constant, a *quadratic* function will fit the table. It’s this phenomenon we want to investigate in this chapter. More precisely:

Let’s hope they are asked to find *a* formula rather than *the* formula. There are, of course, many functions that agree with this table. Later, we’ll see how to classify them (at least the ones given by polynomials).

- We'll see why a constant m th difference means that a table can be matched by a polynomial of degree m .
- Conversely, we'll prove that a polynomial of degree m has constant m th difference.
- We'll come up with a pleasant and efficient method for *finding* a polynomial that matches a table from its difference table.

Actually, we'll look at fitting a polynomial to a table from three points of view:

1. In section 1.1, we'll introduce a method to which we'll return in Chapter 5. It is based on knowing formulas for the sums of powers.
2. In sections 1.2–1.3, we'll develop a method that goes back to Newton; it's based on a careful analysis of "difference tables" (defined below).
3. In section 1.7, we'll develop a technique attributed to Lagrange that generalizes an important theorem in arithmetic (the "Chinese remainder theorem") to polynomials.

So, by the end of the chapter, you'll have a nice collection of methods for fitting functions to data. And you'll have more:

The "important piece" is that the output for a particular input is the starting output plus the sum of all the differences up to that point. If this isn't clear, it will be shortly.

1. The ideas in section 1.1 form one of the important pieces underlying the fundamental theorem of calculus.
2. And, when we return to the method of section 1.1 in Chapter 5, we'll develop a collection of formulas that puts the famous

$$1 + 2 + \cdots + n = \frac{n(n+1)}{2}$$

into a broader perspective.

3. The main method of this chapter (sections 1.2–1.3) will introduce you to the ideas of "discrete calculus." Discrete calculus is the analogue of ordinary calculus where *derivative* is replaced by *difference*. Discrete calculus is useful in all kinds of situations from solving difference equations (the analogue of differential equations) to finding closed forms for recursively defined functions. From a teaching perspective, it shows that many of the habits of mind underlying calculus are independent from notions of limit.
4. In sections 1.3–1.5, we'll see how writing polynomials in different forms can make certain kinds of information transparent. One of our methods will make it relatively simple to write down a polynomial that agrees with a table if, instead of writing it as a sum of monomials, we write it as a sum of other, equally simple, expressions.
5. And, in converting from one form to another, we'll meet some numbers that, like the entries in Pascal's triangle, show up all over mathematics.
6. This idea of subtracting successive numbers in a table, which seems to be ubiquitous among middle and high school students, turns out to be the germ of a general-purpose strategy for analyzing functions.

In the language of linear algebra, we'll choose different *bases* for our polynomials.

After all, differences tell you something about how the function is changing.

Let's get concrete about all this. First, a couple of introductory problems.

Problems

- In a table like the one below, the Δ column is obtained by *subtracting* successive outputs. Confirm that each output can be obtained from the previous output by *adding* it to the adjacent entry in the Δ column.

Input	Output	Δ
0	3	5
1	8	5
2	13	5
3	18	5
4	23	5
5	28	5
6	33	

We'll use this notation throughout the book: Δ (pronounced "delta") stands for "difference." Beside each output, we put in the Δ column the next output minus the current output. Beside 3 (the output for 0) goes $8 - 3 = 5$. Beside 8 (the output for 1) goes $13 - 8 = 5$. Beside 13 (the output for 2) goes $18 - 13 = 5$. All the differences are 5 in this example.

Let's call this the "up and over" property of the table.

- Show that, in an input-output table like the one below, every output is the sum of the output for 0 and the elements in the Δ column up to the number above the output.

We could say "The outputs are the running totals of the differences."

Input	Output	Δ
0	1	-2
1	-1	12
2	11	38
3	49	76
4	125	126
5	251	188
6	439	262
7	701	

Input	Output	Δ
0	1	-2
1	-1	12
2	11	38
3	49	76
4	125	126
5	251	188
6	439	262
7	701	

Input	Output	Δ
0	1	-2
1	-1	12
2	11	38
3	49	76
4	125	126
5	251	188
6	439	262
7	701	

Let's call this the "hockey stick" property of the table. So, every difference table has both the "up and over" and "hockey stick" properties.

1.1 Doing It with Sums

In this section, we'll look at one of the many methods for finding a polynomial function that agrees with a table. Let's return to the table on page 1.

Input	Output
0	3
1	8
2	13
3	18
4	23
5	28
6	33

The “first differences” are constant:

Input	Output	Δ
0	3	5
1	8	5
2	13	5
3	18	5
4	23	5
5	28	5
6	33	

For “constant difference” examples, you could reason like this:

I can get the 8 by taking $3 + 5$. And 13 is $8 + 5 = (3 + 5) + 5$. Ah, I just keep stacking up 5s:

$$\begin{aligned}
 33 &= 28 + 5 \\
 &= (23 + 5) + 5 = 23 + 2 \cdot 5 \\
 &= (18 + 5) + 2 \cdot 5 = 18 + 3 \cdot 5 \\
 &= (13 + 5) + 3 \cdot 5 = 13 + 4 \cdot 5 \\
 &= (8 + 5) + 4 \cdot 5 = 8 + 5 \cdot 5 \\
 &= (3 + 5) + 5 \cdot 5 = 3 + 6 \cdot 5.
 \end{aligned}$$

In general, to find the output at n , I start with 3 and add “ n ” 5s. So, the formula I want is $5n + 3$. A function that agrees with the table is given by $f(n) = 5n + 3$.

This simple example already exhibits many features of the more general picture:

This is the “up and over” feature of difference tables (see problem 1 on page 3).

1. In a table like this, each output can be obtained from the previous output a by adding a and the entry in the Δ column to the right of a .
2. A constant Δ column means that a degree 1 (that is, a *linear*) polynomial will fit the table.
3. Finding a formula that agrees with a table is connected to adding up the numbers in the Δ column. (You can also call the Δ column “the first differences.”)

An application of the hockey stick property (problem 2 on page 3).

Order of operations: $23 + 2 \cdot 5$ means $23 + (2 \cdot 5)$.

Feature 1 is essential to everything we do with difference tables. It's true almost by definition (think of how you *get* the Δ column). You'll handle the details of establishing Feature 2 in the problem set. Feature 3 will come up over and over in this book; we'll deal with it in detail in Chapter 5.

Let's look at another example. We have a mystery function g ; what could a formula for g be?

n	$g(n)$
0	1
1	-2
2	1
3	10
4	25
5	46
6	73

Here's the table with *two* difference columns:

n	$g(n)$	Δ	Δ^2
0	1	-3	6
1	-2	3	6
2	1	9	6
3	10	15	6
4	25	21	6
5	46	27	
6	73		

So, Δ isn't constant. Ah, but *its* first differences *are* constant. Cover up the $g(n)$ column and make believe you were trying to find a formula for the Δ column; its outputs go up by the same amount each time, so there's a linear polynomial that produces the Δ column. In fact, it's $6n - 3$:

n	Δ	$\Delta^2 = \Delta\Delta$
0	-3	6
1	3	6
2	9	6
3	15	6
4	21	6
5	27	

Input	Output	Δ
0	3	5
1	8	5
2	13	5
3	18	5
4	23	5
5	28	5
6	33	

As we'll see later, there are many functions that agree with this table. In some sense, we're looking for the "simplest" function g that tabulates like this.

We're using Δ^2 as shorthand for $\Delta\Delta$, the "second differences" or the "differences of the differences." This use of exponent notation is not very standard in precollege mathematics, but it's pretty common in post-secondary courses. One has to rely on context to know the operation to which the exponent refers; in this case, it's not repeated multiplication but repeated *application* of the Δ process.

One strategy in this game is to keep taking differences until you either run out of outputs or get a constant difference. Notice that the next difference column (the Δ^3 column) will be constant, too; it will be all 0s.

To get the entry in the Δ column for an input of n , start with -3 and add n 6s.

We might be able to make some progress on g now that we are equipped with this formula for the first differences.

n	$g(n)$	$\Delta = 6n - 3$
0	1	$-3 = 6 \cdot 0 - 3$
1	-2	$3 = 6 \cdot 1 - 3$
2	1	$9 = 6 \cdot 2 - 3$
3	10	$15 = 6 \cdot 3 - 3$
4	25	$21 = 6 \cdot 4 - 3$
5	46	$27 = 6 \cdot 5 - 3$
6	73	

Let's see where $g(5)$ comes from, using the hockey stick property, replacing each number in the Δ column by the calculation that led to it:

n	$g(n)$	$\Delta = 6n - 3$
0	1	-3
1	-2	3
2	1	9
3	10	15
4	25	21
5	46	27
6	73	

$$\begin{aligned}
 g(5) &= 1 + \boxed{-3} + \boxed{3} + \boxed{9} + \boxed{15} + \boxed{21} \\
 &= 1 + \boxed{6 \cdot 0 - 3} + \boxed{6 \cdot 1 - 3} + \boxed{6 \cdot 2 - 3} + \boxed{6 \cdot 3 - 3} + \boxed{6 \cdot 4 - 3}
 \end{aligned}$$

Now group "like terms."

$$\begin{aligned}
 g(5) &= 1 + 6(0 + 1 + 2 + 3 + 4) + (-3 + -3 + -3 + -3 + -3) \\
 &= 1 + 6(0 + 1 + 2 + 3 + 4) + 5(-3).
 \end{aligned}$$

Similarly (try it):

$$\begin{aligned}
 g(6) &= 1 + 6(0 + 1 + 2 + 3 + 4 + 5) + (-3 + -3 + -3 + -3 + -3 + -3) \\
 &= 1 + 6(0 + 1 + 2 + 3 + 4 + 5) + 6(-3).
 \end{aligned}$$

Well, at least if $n = 0, \dots, 6$.

And, in general, a function g that agrees with the table will have the property that

$$g(n) = 1 + 6(0 + 1 + 2 + 3 + 4 + 5 + \dots + (n-1)) + n(-3).$$

Now, there's a famous formula for the sum of the first n integers:

$$0 + 1 + 2 + 3 + \cdots + (n - 1) = \frac{n(n - 1)}{2}.$$

If you believe that, we have

$$\begin{aligned} g(n) &= 1 + 6(0 + 1 + 2 + 3 + 4 + 5 + \cdots + (n - 1)) + n \cdot -3 \\ &= 1 + 6\left(\frac{n(n - 1)}{2}\right) - 3n \\ &= 3n^2 - 6n + 1. \end{aligned}$$

If you work through some more examples, it becomes clear that a table with a constant *second* difference can be matched by a polynomial of degree 2.

But proceeding along the lines we've started here will take some work. A constant first difference required us to add up n identical constants. A constant second difference required us to add up the first n integers. It turns out that a constant third difference requires us to add up the first n *squares*. And in general, a constant difference in the m th column requires that you are able to add up the first $n(m - 1)$ st powers. We'll get formulas for such sums in Chapter 5, and we'll return to this approach then, finishing the story. This will give us a perfectly serviceable method for "resolving" difference tables, once we have nice formulas for sums of powers.

You can check that this agrees with the table.

The sum of constants ($m = 0$) is a linear function, the sum of a linear function ($m = 1$) is a quadratic, the sum of squares turns out to be a cubic, and so on. If we knew this for sure, we'd be able to prove that, if the m th differences are constant, there's a polynomial of degree m that fits the table.

Ways to think about it

We started down a natural path of investigation and came up with a promising method for finding a polynomial that agrees with a table: Take the differences until you get a constant, then successively find a formula for each column (by summing the column to the right of it). Trouble is, these formulas get harder to find as the power goes up.

This happens all the time in mathematics research. The solution of one problem is reduced to the solution of another. Sometimes, the second problem is easier to solve, sometimes it's not. When it's not, as in our case, there are two things one can do:

- Try another approach.
- Dig in, pour some coffee, and work on the "reduced" problem.

We'll do both in this book. In the next section, we'll look at a slightly different approach for fitting polynomials to data that is simple, computationally feasible, and fun. But the question of summation formulas is too intriguing to leave, and it has applications beyond finding functions that agree with tables. So, in Chapter 5, we'll return to the topic of summations, and we'll get the formulas we need to finish off the approach we started here.

Have we solved the problem of finding a function that agrees with a table? Well, we've reduced it to doing something else (finding summation formulas). Sometimes, people in this position say they've solved the problem "in principle." If nothing else, that makes you feel better.

Problems

3. Find a polynomial function that agrees with each table.

(a)

Input	Output
0	1
1	4
2	7
3	10
4	13
5	16

(b)

Input	Output
0	3
1	6
2	9
3	12
4	15
5	18

(c)

Input	Output
0	3
1	4
2	8
3	15
4	25
5	38

(d)

Input	Output
0	5
1	6
2	10
3	17
4	27
5	40

(e)

Input	Output
0	5
1	7
2	12
3	20
4	31
5	45

(f)

Input	Output
0	10
1	14
2	24
3	40
4	62
5	90

4. Find a function that agrees with this table:

Input	Output	Δ
0	c	d
1		d
2		d
3		d
4		d
5		

Legend has it that the young Gauss figured out this formula and used it to add up the integers between 1 and 100 in a couple seconds. There's also a rumor that he used no such formula—he just added the numbers in his head.

If you get stuck, you can find proofs in [15] or on the web. You can always use mathematical induction on these things, but try to find a proof that shows how someone might have come upon the formulas.

5. (a) While you are at it, show that the famous formula for the sum of the first n integers:

$$0 + 1 + 2 + 3 + \cdots + (n - 1) = \frac{n(n - 1)}{2}.$$

(b) While you are at it, show that

$$0^2 + 1^2 + 2^2 + 3^2 + \cdots + (n - 1)^2 = \frac{n(n - 1)(2n - 1)}{6}.$$

6. Find a function that agrees with this table:

Input	Output	Δ	Δ^2
0	3	4	2
1			2
2			2
3			2
4			
5			

You might want to fill in the table first. Is there only one table that agrees with these "boundary conditions?"

7. Find a function that agrees with this table:

Input	Output	Δ	Δ^2
0	a	b	c
1			c
2			c
3			c
4			
5			

Recast your solution into the statement of a theorem.

8. Find a possible formula for $f(n)$:

n	$f(n)$
0	1
1	-1
2	11
3	49
4	125
5	251
6	439
7	701

Maybe use the result from problem 5?

9. (a) Suppose you knew that the formula for the sum of the first n cubes is a polynomial in n of degree 4. That is:

$$0^3 + 1^3 + 2^3 + \dots + n^3 = a_4 n^4 + a_3 n^3 + a_2 n^2 + a_1 n + a_0.$$

Use this to prove that, if the 4th differences in a table are constant, a polynomial of degree 4 will fit the table exactly.

(b) Suppose you knew that the formula for the sum of the first n $(m - 1)$ th powers is a polynomial in n of degree m . Use this to prove that, if the m th differences in a table are constant, a polynomial of degree m will fit the table exactly.

It really is true. In Chapter 5, we'll derive a whole sequence of formulas like this:

m	$0^m + 1^m + \dots + (n-1)^m$
0	$n - 1$
1	$\frac{n(n-1)}{2}$
2	$\frac{n(n-1)(2n-1)}{6}$
3	$\frac{n^2(n-1)^2}{4}$