

## 4 Sequences as Functions

We've seen many different types of function in this course and used them to model various situations. In practice, one is often faced with the opposite problem: given experimental data, what type of function should you try?

### 4.1 Polynomial Sequences: First, Second, and Higher Differences

To begin to answer this, first ask yourself, "What is a sequence?" Hopefully you have a decent intuitive idea already. More formally, a sequence is a function whose domain is a set like the natural numbers, for example

$$f : \mathbb{N} \rightarrow \mathbb{R} : n \mapsto 3n^2 - 2$$

defines the sequence

$$(f(1), f(2), f(3), \dots) = (1, 10, 25, 46, 73, \dots)$$

This is indeed the intuitive idea of a function to many grade-school students: continuity and domains including fractions or even irrational numbers are more advanced concepts.

Suppose instead that all you have is a data set

$x$	1	2	3	4	5
$y$	1	10	25	46	73

perhaps arising from an experiment. Could you recover the original function  $y = f(x)$  directly from this data? You could try plotting data points as we've done, though it is hard to decide directly from the plot whether we should try a quadratic model, some other power function/polynomial, or perhaps an exponential. Of course, the physical source of real-world data might also provide clues.

A more mathematical approach involves considering how data values *change*:

$x$	1	$\xrightarrow{+1}$	2	$\xrightarrow{+1}$	3	$\xrightarrow{+1}$	4	$\xrightarrow{+1}$	5
$y$	1	$\xrightarrow{+9}$	10	$\xrightarrow{+15}$	25	$\xrightarrow{+21}$	46	$\xrightarrow{+27}$	73
		$\xrightarrow{+6}$		$\xrightarrow{+6}$		$\xrightarrow{+6}$			

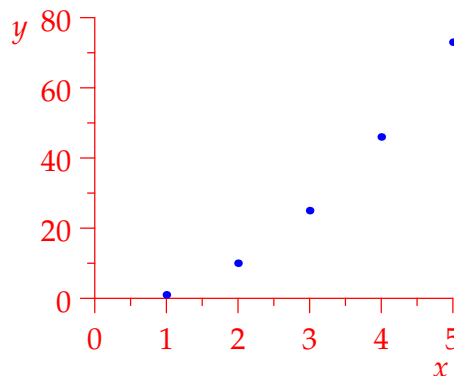
The **first-differences** in the  $x$ -values are constant whereas those for the  $y$ -values are increasing

$$(y_{n+1} - y_n) = (9, 15, 21, 27, \dots)$$

You likely already notice the pattern: the sequence of first-differences is increasing *linearly* as the *arithmetic sequence*

$$y_{n+1} - y_n = 3 + 6n$$

To make this even clearer, note that the sequence of **second-differences** in the  $y$ -values is *constant* (+6). These facts are huge clues that we expect a quadratic function.



But why? Well we can certainly check the following directly:

**Linear Model** If  $f(n) = an + b$ , then the sequence of first-differences is constant

$$f(n+1) - f(n) = a$$

**Quadratic Model** If  $f(n) = an^2 + bn + c$ , then the sequence of first-differences is linear and the second-differences are constant:

$$g(n) := f(n+1) - f(n) = 2an + a + b, \quad g(n+1) - g(n) = 2a$$

The relationship between these results and the *derivative(s)* of the original function  $f(x)$  should feel intuitive: what happens if you differentiate a quadratic twice?

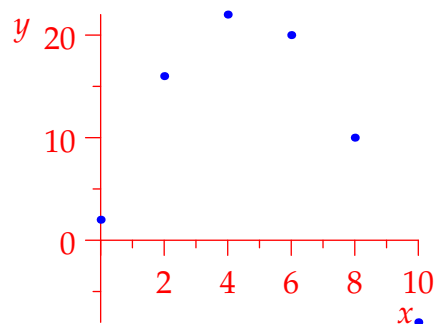
**Example 4.1.** You are given the following data set

$x$	0	2	4	6	8	10
$y$	2	16	22	20	10	-8

The  $x$ -values have constant first-differences while the  $y$ -values have constant second-differences

First-differences: 14, 6, -2, -10, -18

Second-differences: -8, -8, -8, -8



We therefore suspect a quadratic model  $y = f(n) = an^2 + bn + c$ . Rather than using the above formulae, particularly since the  $x$ -differences are not 1, it is easier just to substitute:

$$2 = y(0) = c, \quad \begin{cases} 16 = f(2) = 4a + 2b + 2 \\ 22 = f(4) = 16a + 4b + 2 \end{cases} \implies \begin{cases} 2a + b = 7 \\ 8a + 2b = 10 \end{cases} \implies 4a = -4$$

whence  $a = -1$ ,  $b = 9$  and  $c = 2$ . A quadratic model is therefore

$$y = f(n) = -n^2 + 9n + c = -n^2 + 9n + 2$$

It is easily verified that the remaining data values satisfy this relationship.

There are at least two issues with our method:

1. The question we're answering is, "Find a quadratic model satisfying given data." Constant second-differences don't guarantee that only a quadratic model is suitable. For example,

$$y = -n^2 + 9n + 2 + 297n(n-2)(n-4)(n-6)(n-8)(n-10)$$

is a very complicated model satisfying the same data set!

2. It is very unlikely that experimental data will fit such precise patterns (why not?). However, if the differences are *close* to satisfying such patterns, then you should feel confident that a linear/quadratic model is a good choice.

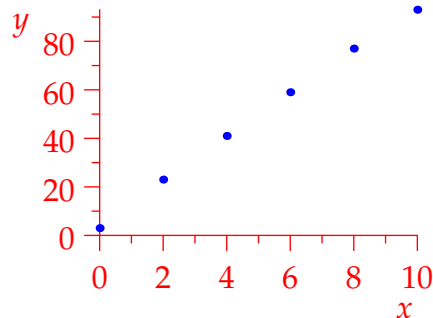
**Example 4.2.** Given the data set

$x$	0	2	4	6	8	10
$y$	3	23	41	59	77	93

with sequences of first- and second-differences

First-differences: 20, 18, 18, 18, 16

Second-differences:  $-2, 0, 0, -2$



do you think a linear or quadratic model would be superior?

If you wanted a linear model, you'd likely be inclined to try  $f(x) = 9x + b$  for some constant  $b$ . Here are two options:

1.  $f(x) = 9x + 5$  fits the middle four data values perfectly, but as a predictor is too large at the endpoints:  $f(0) = 5 > 3$  and  $f(10) = 95 > 93$ .
2.  $f(x) = 9x + 5 - \frac{2}{3}$  doesn't pass through any of the data values but seems to reduce the net error to zero:

$$\begin{array}{c|cccccc} x & 0 & 2 & 4 & 6 & 8 & 10 \\ \hline f(x) - y & -\frac{4}{3} & \frac{2}{3} & \frac{2}{3} & \frac{2}{3} & \frac{2}{3} & -\frac{4}{3} \end{array} \implies \sum_x f(x) - y = 0$$

Neither model is perfect, but then this is what you expect with real-world data!

**Exercises 4.1.** 1. For each data set, find a function  $y = f(x)$  modelling the data.

(a) 

$x$	2	4	6	8
$y$	-1	2	7	14

(b) 

$x$	2	5	8	11	14
$y$	-6	-15	-6	21	66

(c) 

$x$	0	6	9	15
$y$	3	15	21	33

(Be careful with (c): the  $x$ -differences aren't constant!)

2. Suppose a table of data values containing  $(x_0, y_0)$  has constant first-differences in both variables

$$\Delta x = x_{n+1} - x_n = a, \quad \Delta y = b$$

Find the equation of the linear function  $y = f(x)$  through the data.

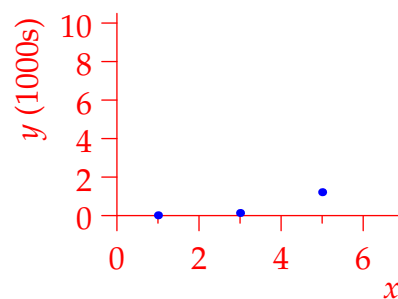
3. What relationship do you expect to find with the sequential differences of a cubic function  $f(n) = an^3 + bn^2 + cn + d$ ? What about a degree- $m$  polynomial  $f(n) = an^m + bn^{m-1} + \dots$ ?
4. If  $f(n) = an^2 + bn + c$  is a quadratic model for the data in Example 4.2 with constant second-differences  $-1$ , show that  $a = -\frac{1}{8}$ . What might be reasonable values for  $b, c$ ?
5. (Hard) Suppose  $f(x)$  is a twice-differentiable function and  $h > 0$  is constant. Use the mean value theorem from calculus to explain the following.
  - (a) First-differences  $f(x+h) - f(x)$  are proportional to  $f'(\xi)$  for some  $\xi \in (x, x+h)$ .
  - (b) Second-differences satisfy  $(f(x+2h) - f(x+h)) - (f(x+h) - f(x)) = f''(\xi)h\alpha$  for some  $\xi$  between  $x$  and  $x+h$  and some  $\alpha$ . Why is it unlikely that  $\alpha$  is constant?

## 4.2 Exponential, Logarithmic & Power Sequences

To observe relationships between data values, you might also have to consider *ratios* between successive terms or skip values.

**Example 4.3.** From a first glance at the given data, it is hard to decide whether an exponential or a quadratic (or higher degree polynomial) model is more suitable. If we try to apply the constant-difference method, we don't seem to get anything helpful:

$x$	1	$\xrightarrow{+2}$	3	$\xrightarrow{+2}$	5	$\xrightarrow{+2}$	7
$y$	15	$\xrightarrow{+120}$	135	$\xrightarrow{+1080}$	1215	$\xrightarrow{+9720}$	10935
		$\xrightarrow{+960}$		$\xrightarrow{+8640}$			



By the time we're looking at second-differences, any conclusion would be very weak since we only have two data values!

If instead we think about *ratios* of  $y$ -values, then a different pattern emerges:

$x$	1	$\xrightarrow{+2}$	3	$\xrightarrow{+2}$	5	$\xrightarrow{+2}$	7
$y$	15	$\xrightarrow{\times 9}$	135	$\xrightarrow{\times 9}$	1215	$\xrightarrow{\times 9}$	10935

The question remains: what type of function scales its output by 9 when 2 is added to its input:  $f(x+2) = 9f(x)$ ? This is a function that converts *addition* to *multiplication*: an exponential! If we try  $y = f(x) = ba^x$  for some constants  $a, b$ , then

$$f(x+2) = ba^{x+2} = ba^2b^x = a^2f(x)$$

from which a suitable model is  $y = 5 \cdot 3^x$ .

We can see the pattern in the example more generally:

**Exponential Model** If  $f(x) = ba^x$ , then adding a constant to  $x$  results in

$$f(x+k) = ba^{x+k} = a^k f(x)$$

If  $x$ -values have constant differences ( $+k$ ), then  $y$ -values will be related by a constant *ratio* ( $\times a^k$ ). You might remember this as 'addition-product' or 'arithmetic-geometric.'

Such a simple pattern is often disguised:

- Complete data might not be given so you might have to skip some data values to see a pattern. For example, if our original data was

$x$	1	3	4	5	7
$y$	15	135	405	1215	10935

then the  $x$ -values are not in a strictly arithmetic sequence.

- As in Example 4.2, real-world/experimental data will only *approximately* exhibit such patterns.

**Example 4.4.** A population of rabbits is measured every two months resulting in the data set

$t$	0	2	4	6	8	10
$P$	5	7	10	14	19	28

The data seems very close to being quadratic; consider the first and second sequences of  $P$ -differences

$$\Delta P = (2, 3, 4, 5, 9), \quad \Delta\Delta P = (1, 1, 1, 4)$$

However, the last difference doesn't fit the pattern. Instead, the fact that we expect an exponential model is buried in the experiment: the data is measuring population growth! We therefore instead consider the ratios of  $P$ -values:

$t$	0	2	4	6	8	10
$P$	5	7	10	14	19	28
		$\times 1.4$	$\times 1.43$	$\times 1.4$	$\times 1.36$	$\times 1.47$

The ratios are very close to being constant, whence an exponential model is suggested! To exactly match the first and last data values, we could take the model

$$P(t) \approx 5 \left( \frac{28}{5} \right)^{\frac{t}{10}}$$

$t$	0	2	4	6	8	10
$P$	5	7.057	9.960	14.057	19.839	28

Only  $P(19)$  doesn't match when we take rounding to the nearest integer into account.

We've seen that addition-addition corresponds to a linear model and that addition-multiplication to an exponential. There are two other natural combinations.

**Logarithms** These operate exactly as exponentials but in reverse. If  $f(n) = \log_a x + b$ , then *multiplying*  $x$  by a constant results in a constant *addition/subtraction* to  $y$ :

$$f(kx) = \log_a(kx) + b = \log_a k + \log_a x + b = \log_a k + f(x)$$

This could be summarized as 'product-addition.'

**Power Functions** If  $f(x) = ax^m$ , then multiplying  $x$  by a constant will do the same to  $y$

$$f(kx) = a(kx)^m = ak^m x^m = k^m f(x)$$

We have a 'product-product' relationship between successive terms.

**Examples 4.5.** Find the patterns in the following data and suggest a model  $y = f(x)$  in each case.

$x$	6	18	54	162
$y$	1	2	3	4

$x$	3	6	9	12
$y$	135	1080	3645	8640

The sequential approach in this chapter is a form of *discrete calculus*: using a pattern of *differences* to predict the original function is similar to how we use knowledge of a derivative  $f'(x)$  to find  $f(x)$ .

**Example 4.6.** Suppose  $g(2) = 3$  and  $g(4) = 9$ . What do you think should be the value of  $g(8)$ ?

It depends on the type of model you try.

1. For a linear (addition-addition) model we know that  $\Delta x = 2$  corresponds to  $\Delta y = 6$ , so

$$g(8) = g(4 + 2\Delta x) = g(4) + 2\Delta y = 9 + 12 = 21$$

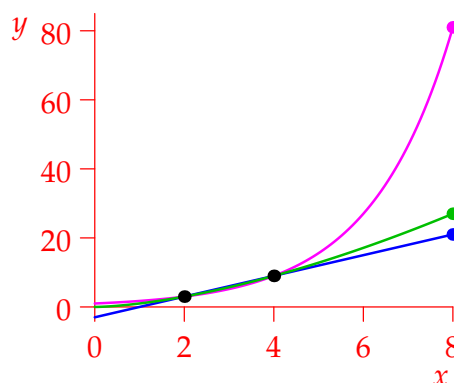
2. For an exponential (addition-product) model,  $\Delta x = 2$  corresponds to a  $y$ -ratio  $r_y = \frac{9}{3} = 3$ , so

$$g(8) = r_y g(6) = r_y^2 g(4) = 9 \cdot 9 = 81$$

3. For a power (product-product) model,  $r_x = 2$  corresponds to a  $r_y = 3$ , so

$$g(8) = g(2 \cdot 4) = g(4r_x) = r_y g(4) = 3 \cdot 9 = 27$$

We do not need to calculate the models explicitly(!), though they are stated below the graph for convenience.



1.  $g(x) = 3x - 3$

2.  $g(x) = 3^{x/2}$

3.  $g(x) = x^{\log_2 3}$

**Exercises 4.2.** 1. Find the patterns in the following data sets and use them to find a model  $y = f(x)$ .

(a)	$x$	0	1	2	3	4
	$y$	80	120	180	270	405

(b)	$x$	2	4	8	10
	$y$	1	16	256	625

(c)	$x$	1	3	5	7	9
	$y$	15	5	19	57	119

(d)	$x$	1	3	4	6
	$y$	1	36	216	7776

(e)	$x$	20	60	180	540
	$y$	2	4	6	8

(f)	$x$	2	6	54	486	4374
	$y$	2	4	8	12	16

2. Take logarithms of the power relationship  $y = ax^m$ . What is the relationship between  $\ln y$  and  $\ln x$ ? Use this to give another reason why the inputs and outputs of power functions satisfy a 'product-product' relationship.
3. How does our analysis of exponential functions change if we add a constant to the model? That is, how might you recognize a sequence arising from a function  $f(x) = ba^x + c$ ?
4. Suppose  $f(5) = 12$  and  $f(10) = 18$ . Find the value of  $f(20)$  supposing  $f(x)$  is a:
  - (a) Linear function;
  - (b) Exponential function;
  - (c) Power function.

If  $f(20) = 39$ , which of the three *models* do you think would be more appropriate?

### 4.3 Newton's Method

To finish our discussion of sequences we revisit a (hopefully) familiar technique for approximating solutions to equations. Newton's method is particularly useful for polynomials since the computations produce sequences of *rational numbers*. Variations of this approach have been in use for thousands of years.

**Example 4.7.** We motivate the method by considering an ancient method for approximating  $\sqrt{2}$ , known to the Babylonians 2500 years ago!

Suppose  $x_n > \sqrt{2}$ . Then  $\frac{2}{x_n} < \frac{2}{\sqrt{2}} = \sqrt{2}$ . It seems reasonable to guess that their average

$$x_{n+1} = \frac{1}{2} \left( x_n + \frac{2}{x_n} \right)$$

should be a more accurate approximation to  $\sqrt{2}$ . If start with an initial guess  $x_0 = 2$ , then we obtain the sequence

$$x_1 = \frac{1}{2} \left( 2 + \frac{1}{2} \right) = \frac{3}{2}, \quad x_2 = \frac{17}{12} = 1.4166\dots, \quad x_3 = \frac{577}{408} = 1.4142\dots, \quad \dots$$

This sequence certainly *appears* to be converging to  $\sqrt{2}$ ...

This approach is sometimes called the *method of the mean*. It may be applied to any square-root  $\sqrt{a}$  where  $a > 0$ : let  $x_0 > 0$  and define,

$$x_{n+1} := \frac{1}{2} \left( x_n + \frac{a}{x_n} \right) \tag{*}$$

A rigorous proof that the sequence converges requires more detail than is appropriate for us, but two observations should make it seem more believable:

1. If the sequence (\*) has a limit  $L$ , then it must satisfy

$$L = \frac{1}{2} \left( L + \frac{a}{L} \right) \implies 2L^2 = L^2 + a \implies L^2 = a \implies L = \sqrt{a}$$

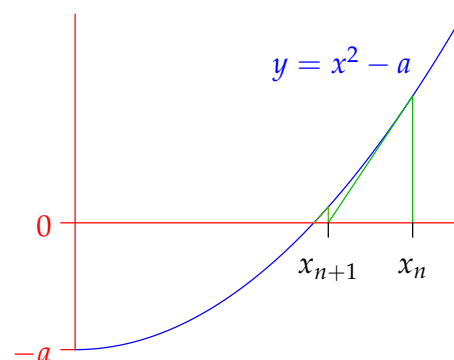
where we take the positive root since all terms  $x_n$  are plainly positive.

2. The iterations have a convincing *geometric* interpretation: the sequence of iterates can be found by repeatedly taking the **tangent line** to the **curve**  $y = f(x) = x^2 - a$  and intersecting it with the  $x$ -axis. To see why, note that the tangent line at  $x_n$  has equation

$$y = f(x_n) + f'(x_n)(x - x_n) = x_n^2 - a + 2x_n(x - x_n)$$

which intersects the  $x$ -axis ( $y = 0$ ) when

$$x = x_n + \frac{a - x_n^2}{2x_n} = \frac{1}{2} \left( x_n + \frac{a}{x_n} \right) = x_{n+1}$$

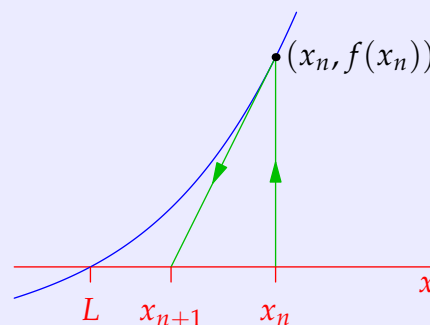


The geometric idea generalizes...

**Definition 4.8.** Given a differentiable function  $f(x)$  with non-zero derivative, the *Newton–Raphson iterates* of an initial value  $x_0$  are defined by

$$x_{n+1} := x_n - \frac{f(x_n)}{f'(x_n)}$$

If  $L = \lim_{n \rightarrow \infty} x_n$  exists, then it is a root of the function:  $f(L) = 0$ .



The picture explains why we might expect this to work. The **tangent line** to the graph at the point  $(x_n, f(x_n))$  has equation

$$y = f(x_n) + f'(x_n)(x - x_n)$$

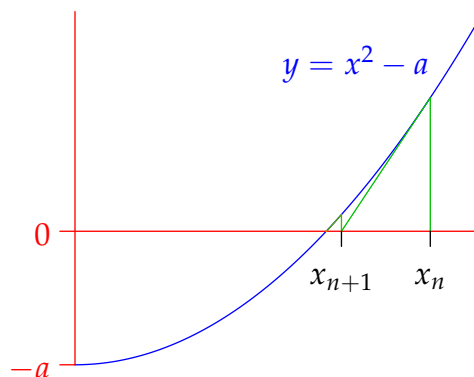
and thus  $x$ -intercept ( $y = 0$ ) at  $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$ .

**Example 4.9.** To approximate  $\sqrt{2}$ , consider  $f(x) = x^2 - 2$  and a starting value of  $x_0 = 1$  since this is reasonably close to  $\sqrt{2}$ . Now iterate

$$x_{n+1} = x_n - \frac{x_n^2 - 2}{2x_n} = \frac{x_n^2 + 2}{2x_n}$$

which yields the sequence (to 4 d.p.)

$$\left(1, \frac{3}{2}, \frac{17}{12}, \frac{577}{408}, \dots\right) = (1, 1.5, 1.4166\dots, 1.4142\dots)$$



This approach to approximating square roots was known to the ancient Babylonians 2500 years ago! For approximating the square root  $\sqrt{a}$  it is sometimes known as the *method of the mean*, since

$$x_{n+1} = \frac{1}{2} \left( x_n + \frac{a}{x_n} \right)$$

is the average of  $x_n$  and the  $\frac{a}{x_n}$ . The rough (pre-calculus) idea is that if  $x_n > \sqrt{a}$ , then  $\frac{a}{x_n} < \sqrt{a}$  so their average should be in between.



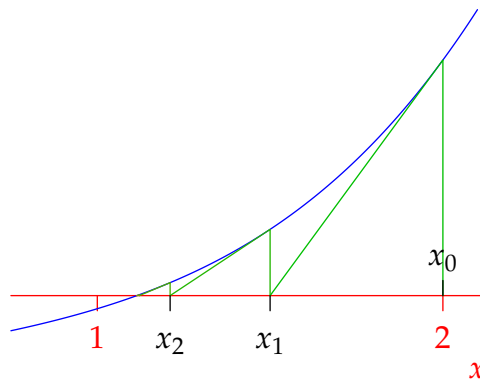
**Example 4.10.** To find a root of  $f(x) = x^4 + 4x - 6$ , start with  $x_0 = 2$  and iterate

$$x_{n+1} = x_n - \frac{x_n^4 + 4x_n - 6}{4x_n^3 + 4} = \frac{3(x_n^4 + 2)}{4(x_n^3 + 1)}$$

which yields the sequence (to 4 d.p.)

$$\left(2, \frac{3}{2}, \frac{339}{280}, \dots\right) = (2, 1.5, 1.2107, 1.1214, 1.1144, 1.1144, \dots)$$

Since the sequence seems to be converging, we conclude that 1.1144 is approximately a root.



This method can be attempted for any differentiable function, though the sequence isn't guaranteed to converge: play with some examples such as  $f(x) = x^2 - 2$ , and  $f(x) = x^3 - 5x$ . You can find graphical interfaces online for this (for instance with Geogebra).