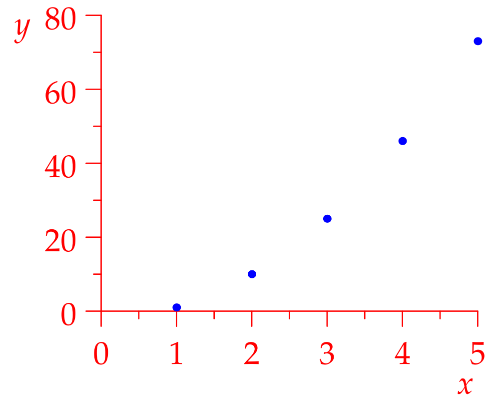
5 Sequences as Functions

Often one is faced with data values to which one wants to fit a function. What type of function should you try?

To begin to answer this, first ask yourself, "What is a sequence?" Hopefully you have a decent working definition handy. In a formal sense, a sequence is simply a function whose domain is some countable set, for instance

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



defines the sequence

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

Suppose now that all you were given was a data set

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

Could you recover the above function $y = f(x)$ from this data? You could try several things, including plotting the data points as we've done above. It is hard to look at this and tell whether we are looking at a quadratic model, or some other power function, or perhaps an exponential. Of course the source of the data might also give you some clues.

Another approach is to think about how the data values *change* as we move along the table.

$$\begin{array}{ccccccccccc}
x & & 1 & \xrightarrow{+1} & 2 & \xrightarrow{+1} & 3 & \xrightarrow{+1} & 4 & \xrightarrow{+1} & 5 \\
y & & 1 & \xrightarrow[+9]{} & 10 & \xrightarrow[+15]{} & 25 & \xrightarrow[+21]{} & 46 & \xrightarrow[+27]{} & 73 \\
 & & & & \xrightarrow[+6]{} & & \xrightarrow[+6]{} & & \xrightarrow[+6]{} & &
\end{array}$$

In this case the first differences in the x values are constant whereas those for the y values are increasing

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

It is when we consider the *second differences* in the y values, that is the differences between the terms of the sequence of first differences $(*)$, that we see a constant change. This is a huge clue that we expect a quadratic function.
Indeed the following should be very easy to verify:

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

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

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

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

Exercises

1. If you start only with a data set such as that above (and *not* a function $f(x)$), why can you not be certain that you have a quadratic model? That is, why do the implications just mentioned only go in one direction?

2. What is the relationship between these results and the *derivative(s)* of the original function $f(x)$?

3. What do you think will happen if you have a cubic function $f(n) = an^3 + bn^2 + cn + d$? A degree m polynomial $f(n) = an^m + bn^{m-1} + \cdots$? Can you prove any of your assertions?

4. If your x values have constant first differences, but not necessarily 1, can we make the same conclusions?

Example 5.1. You are given the following data set

x	0	2	4	6	8	10
y	0	14	20	18	8	-10

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

- 1^{st} differences: $14, 6, -2, -10, -18$
- 2^{nd} differences: $-8, -8, -8, -8$

We therefore suspect a quadratic model $y = f(n) = an^2 + bn + c$ where $2a$ would be the constant 2^{nd} differences *if* the x-values were separated by 1. Since the x-values are separated by 2, we instead have

$$2a = -2 \implies a = -1$$

The first differences, should have the form

$$f(n+2) - f(n) = 4an + 4a + 2b = -4n - 4 + 2b$$

which fits the pattern of decreasing by 8 each time. Indeed, with $n = 0$ we quickly see that

$$14 = -4 + 2b \implies b = 9 \implies f(n) = -n^2 + 9n + c$$

It is then a trivial matter to verify that $c = 0$.

With a real data set, arising from a real experiment, it is very unlikely that the data will fit such a precise pattern. If the differences are *close* to satisfying such patterns however, then you should feel confident in searching for a linear or quadratic model.

Exercise Given the data set

x	0	2	4	6	8	10
y	3	22	38	52	64	73

do you think a linear or quadratic model is likely superior? What happens if you ask a spreadsheet?

More generally, you might have to look for other relationships between the data values. Perhaps you might be *multiplying* by a constant to obtain the next term, or perhaps you have to skip some terms to spot easily spot a pattern.

Exercises For each of the following data sets, try to describe the relationships between successive x and y values:

1.

x	2	4	6	8	10
y	4	9	14	19	24

2.

x	1	3	5	7	9
y	15	5	19	57	119

3.

x	1	3	5	7
y	15	135	1215	10935

4.

x	3	6	9	12
y	135	1080	3645	8640

5.

x	6	18	54	162
y	1	2	3	4

At least the first two of these should seem familiar. For the others, consider other types of functions that we've covered, namely power, exponential and logarithmic functions. What sort of relationships between successive x and y values do you expect to see in such situations?

Try the above exercise *before* reading what follows! Just as previously with polynomial functions, we can consider how successive terms are related for power, exponential and logarithmic functions.

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

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

If the x values have constant differences, the y values will be related by a constant *ratio.* Alternatively, you might try to remember this as "addition–product" or "arithmetic–geometric."

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; i.e. we have a "product–product" relationship between successive terms.

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

Exercises 1. By taking logarithms of the power relationship $y = ax^m$, why is it obvious that we should have a "product–product" relationship between successive data values?

2. 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?