AS1056 - Mathematics for Actuarial Science. Chapter 2, Tutorial 2.

Emilio Luis Sáenz Guillén

Bayes Business School. City, University of London.

October 27, 2023

Big-O and little-o notation

Big-O notation

- **Purpose:** Describes an upper bound on the time complexity of an algorithm in terms of the worst-case scenario.
- **Usage:** Commonly used in computer science to analyse the efficiency of algorithms.
- **Example:** If an algorithm has a time complexity of $O(n^2)$, it means that in the worst case, the number of operations grows quadratically with the size of the input.

Big-O notation

Definition (I)

Let f and g be functions from $\mathbb{R}$ to $\mathbb{R}$. We say that,

$$f(x) = O(g(x)) \text{ as } x \to \infty$$

if there is **at least one** choice of a constant $M > 0$, for which you can find a constant k such that:

$$|f(x)| \leq M|g(x)| \quad \text{i.e.} \quad \left|\frac{f(x)}{g(x)}\right| \leq M$$

whenever $x > k$. Beyond some point k, function $f(x)$ is at most a constant M times $g(x)$.

$\longrightarrow f = O(g)$ (big-oh) if eventually f grows slower than some multiple of g

We can also use this notation to describe the behaviour of a function nearby **some** real number a (often $a = 0$).

Definition (II)

We say that,

$$f(x) = O(g(x)) \text{ as } x \to a$$

if there is **at least one** constant M such that,

$$\left| \frac{f(x)}{g(x)} \right| \leq M$$

for sufficiently small x.

The intuition behind big-oh notation is that f is $O(g)$ if $g(x)$ grows as fast or faster than $f(x)$ as $x \to a$.

Big-O and little-o notation

Little-o notation

- **Purpose:** Describes an upper bound, but in a stronger sense than big-O. It indicates that a function grows strictly slower than the comparison function.
- **Usage:** Less common than Big O, but used when we need to express that one function grows strictly slower than another.
- **Example:** If $f(n) = n$ and $g(n) = n^2$ then $f(n) = o(g(n))$ as $n \to \infty$ because $f(n)$ grows strictly slower than $g(n)$.

While big-O gives an upper limit, little-o indicates that the function grows strictly slower than the comparison function.

Little-o notation

Definition (I)

Let f and g be functions from $\mathbb{R}$ to $\mathbb{R}$. We say that,

$$f(x) = o(g(x)) \text{ as } x \to \infty$$

if **for every** constant $M > 0$, there exists a constant k such that whenever $x > k$:

$$|f(x)| < M|g(x)| \quad \text{i.e.} \quad \lim_{x \to \infty} \left| \frac{f(x)}{g(x)} \right| = 0$$

$\longrightarrow f = o(g)$ (little-oh) if eventually f grows slower than **any** multiple of g

Similarly, to describe the behaviour of a function near some real number a (often $a = 0$):

Definition (II)

We say that,

$$f(x) = o(g(x)) \text{ as } x \to a$$

if and only if:

$$\lim_{x \to a} \left| \frac{f(x)}{g(x)} \right| = 0$$

The intuition behind little-oh notation is that f is $o(g)$ if $g(x)$ grows strictly faster than $f(x)$ as x approaches 0.

For the upcoming exercise recall the following proposition from your lecture notes:

Proposition 2.1

The following two statements are equivalent:

1. If f is differentiable at x_0 with derivative $f'(x_0)$
2. As $h \to 0$, $f(x_0 + h) = f(x_0) + hf'(x_0) + o(h)$

And using big-O notation the latter can be expressed as $f(x_0 + h) = f(x_0) + O(h)$.

Exercise 2.8

Use O and o notation to describe the behaviour of the following functions as x approaches the given values:

(ii) $f_2(x) = \sqrt{1 + x^2}$ as $x \to 0$

Goal: understand/describe the behaviour of $f_2(x)$ as x approximates 0.

Little-o notation

Note that:

1. **Function Value at 0**: $f_2(0) = \sqrt{1 + 0^2} = 1$
2. **First Derivative at 0**:

$$f_2'(x) = \frac{1}{2}(1 + x^2)^{-1/2} \times 2x; \quad f_2'(0) = 0$$

$$\underbrace{\hspace{5cm}}$$

$\longrightarrow \boldsymbol{f_2(x) = 1 + o(x)}$ as $\boldsymbol{x \to 0}$

- $f_2'(0) = 0$, i.e., the rate of change of $f_2(x)$ at $x = 0$ is zero. This aligns with the assertion that $f_2(x) = 1 + o(x)$, suggesting that near $x = 0$, the function approaches the constant value 1 more slowly than x and does not increase/decrease linearly with x. Instead, its change is almost negligible compared to a linear rate, indeed it is getting flat/constant.
- The notation $1 + o(x)$ captures this idea: as $x \to 0$, whatever change happens in $f_2(x)$ from the value 1 is significantly lesser than the change in x itself. In other words, $o(x)$ goes faster to 0 than x.

In summary, "$f_2(x) = 1 + o(x)$ as $x \to 0$" reflects that as x gets closer and closer to 0, the function $f_2(x)$ gets closer to 1, and the deviation of $f_2(x)$ from 1 grows at a rate that is slower than the rate at which x approaches 0.

Note: $f_2(x) = 1 + o(x)$ as $x \to 0$ implies that I could also express $f_2(x)$ as $f_2(x) = 1 + O(x)$ as $x \to 0$...

1. $f_2(x) = 1 + o(x)$ as $x \to 0$
   - This indicates that as x approaches 0, the difference between $f_2(x)$ and 1 becomes negligible compared to x. The function $f_2(x) - 1$ grows much slower than x.
2. $f_2(x) = 1 + O(x)$ as $x \to 0$
   - This indicates that as x gets close to 0, the function $f_2(x)$ is close to 1, and any deviation from 1 is at most linear in magnitude with respect to x.

However, it is much more precise and informative to say $f_2(x) = 1 + o(x)$.

Big-O notation

Binomial Theorem for Fractional Exponent

Let $\alpha = \frac{p}{q}$ be a rational number (p, q integers). Then:

$$(1+x)^{\alpha} = 1 + \alpha x + \frac{(\alpha)(\alpha-1)}{2!}x^2 + ... = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k$$

Therefore,

$$f_2(x) = \sqrt{1+x^2} = 1 + \frac{1}{2}x^2 - \frac{1}{8}x^4...$$

The first non-constant term in the expansion of $f_2(x)$ around $x = 0$ is proportional to x^2.

$\longrightarrow \boldsymbol{f_2(x) = 1 + O(x^2)}$ as $\boldsymbol{x \to 0}$

When you're close to 0, the behaviour of $f_2(x)$ differs from the constant function 1 by an amount that is <u>at most</u> proportional to x^2.
Of course we can also say:

$$f_2(x) = 1 + \frac{1}{2}x^2 + O(x^4)$$
$$f_2(x) = 1 + \frac{1}{2}x^2 - \frac{1}{8}x^4 + O(x^6)$$
$$\vdots$$

$f_2(x) = 1 + O(x^2)$ is more informative than $f_2(x) = 1 + O(x)$

Since x^2 grows slower than x near 0, this implies that $f_2(x)$ is even closer to 1 than what is suggested by the $O(x)$ notation. This function behaves like a parabola near 0, which is flatter than a line when close to 0.

(iii) $f_3(x) = \frac{x}{x^2+1}$ as $x \to 3$

- $f_3(3+h) = 0.3 - 0.08h + o(h)$
  - $f_3(3) = 0.3$
  - As x deviates from 3 by a small amount h, the function's value decreases at a rate of 0.08 times that deviation.
  - The term $o(h)$ represents error terms that become negligible compared to h as h approaches 0.

  Our approximation is mostly driven by the 0.3 and the $-0.08h$ components, especially when h is very small.
- $f_3(3+h) = 0.3 + O(h)$
  - $f_3(3) = 0.3$
  - The $O(h)$ notation suggests that the deviation of $f_3(3+h)$ from 0.3 is at most linear in h as h approaches 0.

  However, this notation doesn't specify the exact behaviour or rate of this deviation. It does not specify the exact coefficient in front of h as in the precise derivative calculation. It's a more general representation.

(iv) $f_4(x) = \frac{x-8}{(x+2)(2x-1)}$ as $x \to \frac{1}{2}$

Remember we're interested on the behaviour of $f_4(x)$ as x approaches $\frac{1}{2}$.
Thus consider:

$$f_4\left(\frac{1}{2}+h\right) = \frac{\frac{1}{2}+h-8}{\left(\frac{1}{2}+h+2\right)\left(2\times(\frac{1}{2}+h)-1\right)} =$$

$$= \underbrace{\frac{1}{2h}}_{\to\infty \text{ as } h\to 0} \times \underbrace{\frac{h-7.5}{h+2.5}}_{<\infty \text{ as } h\to 0}$$