A note on o and O

Ulrik Skre Fjordholm

April 15, 2020

1 Little-o and big-O notation

When studying the behavior of a function f, it is often less important exactly what the function looks like near some point $x = a$, than having some rough estimate of its qualitative behavior. The o and O notation, part of the so-called Bachmann–Landau notation, are two ways of quantifying *asymptotic behavior* of $f(x)$ as $x \to a$.

Definition. Let X, Y be normed vector spaces, let $a \in X$, and let $f : X \to Y$ and $\alpha : X \to [0, \infty)$ be functions. When we write

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

we mean that there exists some constants $C > 0$ and $r > 0$ such that

$$\|f(x)\|_Y \leqslant C\alpha(x) \qquad \forall\, x \in B(a; r). \tag{1}$$

When we write

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

we mean that

$$\lim_{x \to a} \frac{\|f(x)\|_X}{\alpha(x)} = 0,$$

that is, for every $\varepsilon > 0$ there exists some $\delta > 0$ such that

$$\frac{\|f(x)\|_Y}{\alpha(x)} \leqslant \varepsilon \qquad \forall\, x \in B(a; \delta). \tag{2}$$

In a similar way, when we write

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

we mean that there exists some $C > 0$ and $K > 0$ such that $\|f(x)\|_Y \leqslant C\alpha(x)$ for all $x \in X$ with $\|x\|_X > K$, and when we write

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

we mean that $\lim_{\|x\| \to \infty} \frac{\|f(x)\|_X}{\alpha(x)} = 0$, that is, for every $\varepsilon > 0$ there exists some $K > 0$ such that $\frac{\|f(x)\|_Y}{\alpha(x)} \leqslant \varepsilon$ for all $x \in X$ with $\|x\|_X > K$.

Remarks.

1. We can rewrite (2) as $\|f(x)\|_Y \leqslant \varepsilon\alpha(x)$ for $x \in B(a; \delta)$. Comparing to (1), we see that the constant C "becomes smaller and smaller" as x approaches a.

2. Note that o and O do not refer to specific functions, but instead convey a property. Therefore, the normal rules of arithmetic do not apply to o and O: For instance, for $x \in \mathbb{R}$ we have both $x = O(|x|)$ and $x^2 = O(|x|)$ as $x \to 0$ (set e.g. $C = 1$, $r = 1$ in the definition), but subtracting these two expressions will not "cancel the O", since $x - x^2 = O(|x|)$ (set e.g. $C = 2$, $r = 1$ in the definition).

3. The o and O expressions are often used casually as if they referred to specific functions; for instance, the expression $f(x) = g(x) + o(\alpha(x))$ as $x \to a$ means $f(x) - g(x) = o(\alpha(x))$ as $x \to a$.

Examples. (i) If $f(x) = (x-a)^2$ for $a, x \in \mathbb{R}$ then $f(x) = o(|x-a|)$ as $x \to a$.

(ii) The number of floating point operations required to multiply an $n \times n$ matrix and an n-vector is $O(n^2)$ as $n \to \infty$. (The exact number is n^2 multiplications and $n(n-1)$ additions, which sums to $2n^2 - n = O(n^2)$ as $n \to \infty$.)

(iii) The function $f(x) = \log x$ (defined for $x \in (0, \infty)$) is $o(x^\alpha)$ as $x \to 0$ for every $\alpha > 0$.

(iv) The expression $f(x) = O(1)$ as $x \to a$ is equivalent to stating that f is locally bounded near a, that is, there are constants $r, C > 0$ such that $\|f(x)\| \leqslant C$ for all $x \in B(x; r)$

(v) The expression $f(x) = o(1)$ as $x \to a$ is equivalent to stating that $f(x) \to 0$ as $x \to a$.

From the definition of O it follows that if $f(x) = O(\alpha(x))$ and $g(x) = O(\beta(x))$ as $x \to a$, then $f(x) + g(x) = O(\alpha(x) + \beta(x))$ as $x \to a$. Some further computational rules are listed below.

Exercise. Prove the following. Each statement is assumed to hold in some limit $x \to a$.

(i) If $f(x) = O(\alpha(x))$ and $g(x) = O(\beta(x))$ then $f(x) + g(x) = O(\alpha(x) + \beta(x))$

(ii) If $f(x) = o(\alpha(x))$ and $g(x) = o(\beta(x))$ then $f(x) + g(x) = o(\alpha(x) + \beta(x))$

(iii) If $f(x) = o(\alpha(x))$ and $g(x) = O(\alpha(x))$ then $f(x) + g(x) = O(\alpha(x))$

(iv) If $f(x) = o(\alpha(x))$ then $f(x) = O(\alpha(x))$

(v) If $f(x) = o(\alpha(x))$ and $g : \mathbb{R} \to \mathbb{R}$ is such that $g(x) = O(\beta(x))$ then $g(x)f(x) = o(\alpha(x)\beta(x))$

(vi) If $f(x) = O(\alpha(x))$ and $g : \mathbb{R} \to \mathbb{R}$ is such that $g(x) = O(\beta(x))$ then $g(x)f(x) = O(\alpha(x)\beta(x))$

2 Little-o and the derivative

Recall the definition of the Fréchet derivative:

Definition. Let X, Y be normed vector spaces, let $O \subset X$ be open and let $F : O \to Y$ be a function. We say that F is *Fréchet differentiable* at a point $a \in O$ if there exists a bounded linear operator $A \in \mathcal{L}(X, Y)$ such that

$$F(x) = F(a) + A(x-a) + o(\|x-a\|) \qquad \text{as } x \to a. \tag{3}$$

We denote the *Fréchet derivative* as $F'(a) = A$. (It is straightforward to prove that the Fréchet derivative is unique, if it exists.)

We will not go into detail about the properties of the Fréchet derivative here, but we will give to different motivations for why the definition looks the way it does.

Recall from Calculus that a function $f : \mathbb{R} \to \mathbb{R}$ is differentiable at $a \in \mathbb{R}$ if the limit

$$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a} \tag{4}$$

exists. This definition does not extend directly to functions between vector spaces $f : X \to Y$ since division by vectors is not a well-defined operation. However, we can rewrite (4) as follows:

$$\begin{aligned}
& & & f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a} & & \\
&\iff & & \lim_{x \to a} \left(\frac{f(x) - f(a)}{x - a} - f'(a) \right) = 0 & & \\
&\iff & & \frac{f(x) - f(a)}{x - a} - f'(a) = o(1) & & \text{as } x \to a \\
&\iff & & f(x) - f(a) - f'(a)(x - a) = o(|x - a|) & & \text{as } x \to a
\end{aligned}$$

(the third and fourth steps following from Example (v) and Exercise (v), respectively). The above expression is almost, but not exactly equal to (3). However, recall that every linear function $A : \mathbb{R} \to \mathbb{R}$ can be written as

$$A(x) = \alpha x$$

for some $\alpha \in \mathbb{R}$. Therefore, if we define the linear operator $A(r) = f'(a)r$ we see that f is differentiable at a (in the Calculus sense) *if and only if* (3) holds.

The second motivation we will consider is in terms of the first-order Taylor expansion with remainder for a function $f : \mathbb{R} \to \mathbb{R}$,

$$f(x) = f(a) + f'(a)(x - a) + \frac{f''(c)}{2}(x - a)^2$$

for some c between a and x. If f is C^2 (i.e., twice continuously differentiable) then we can say that $f''(c)$ is bounded for all c close to a, and therefore

$$f(x) = f(a) + f'(a)(x - a) + O(|x - a|^2) \qquad \text{as } x \to a.$$

Since any term which is $O(|x - a|^2)$ is also $o(|x - a|)$ as $x \to a$, the above implies the definition (3) of the Fréchet derivative. So why do we write $o(|x - a|)$ and not $O(|x - a|^2)$ in (3)? The computation we just performed is valid for all C^2 functions, but there are lots of functions that are not C^2 but still have a reasonable derivative. Consider for instance the function $f(x) = x^\alpha$ for some $\alpha \in (1, 2)$, which is not C^2 near $x = 0$. The tangent at $x = 0$ has slope 0, so it is reasonable that $f'(0) = 0$. However,

$$f(x) - f(0) - 0 \cdot (x - 0) = x^\alpha$$

which is certainly $o(|x|)$, but *not* $O(|x|^2)$ as $x \to 0$.

Last, why do we use an o and not an O in (3)? Consider the statement

$$F(x) = F(a) + A(x - a) + O(\|x - a\|_X) \qquad \text{as } x \to a \tag{5}$$

for some $A \in \mathcal{L}(X, Y)$. Since A is bounded, we have $\|A(x-a)\|_Y \leqslant \|A\|_{\mathcal{L}} \|x-a\|_X$. Thus, $A(x-a) = O(\|x-a\|_X)$ as $x \to a$, so (5) implies that

$$F(x) = F(a) + O(\|x-a\|_X) \qquad \text{as } x \to a.$$

Thus, there exists some $C > 0$ such that $\|F(x) - F(a)\|_Y \leqslant C\|x-a\|_X$ for x close to a, that is, F is "locally Lipschitz". Now note that, if F is any such "locally Lipschitz" function, we can let A be *any* bounded linear operator, since (5) will be true regardless of A. We are lead to the conclusion that if the definition (3) of Fréchet differentiability were replaced by (5), then the Fréchet derivative A of F would not be unique, but could be any bounded linear operator.