Lecture 5

We have figured out that the reasonable definition of the derivative of a map $f : \mathcal{O}_n \to \mathbb{R}^m$ at a point $a \in \mathcal{O}_n$ is $Df(a) = L$ where

$$\lim_{h \to 0} \frac{f(a+h) - f(a) - Lh}{\|h\|} = 0$$

We now proceed to perform a series of reductions that will actually answer the question of differentiability of functions and also provide techniques to compute the derivative.

5.1 Matrix representation of the derivative

Theorem 5.1.1

Let $f : \mathcal{O}_n \to \mathbb{R}^m$ and $a \in \mathcal{O}_n$. Then, f is differentiable at a iff f_i is differentiable at a for all i. In that case, we have

$$Df(a) = \begin{pmatrix} Df_1(a) \\ Df_2(a) \\ \vdots \\ Df_m(a) \end{pmatrix}$$

Proof. Assume that f is differentiable at a and let $L = Df(a)$. For all $i = 1, 2, \ldots, m$ we set $L_i = \Pi_i \circ L$. Further, let $\tilde{f}_i(h) = f_i(a+h) - f_i(a) - L_i h$, so that

$$f(a+h) - f(a) - Lh = (\tilde{f}_1(h), \tilde{f}_2(h), \ldots, \tilde{f}_m(h))$$

But then, $\left|\tilde{f}_i(h)\right| \leq \|f(a+h) - f(a) - Lh\|$ which proves that f_i is differentiable at a and has derivative $Df_i(a) = L_i$, for all i.

Assume that f_i is differentiable at a for all i and define

$$L = \begin{pmatrix} Df_1(a) \\ Df_2(a) \\ \vdots \\ Df_m(a) \end{pmatrix}$$

Let $\Pi_i(f(a+h) - f(a) - Lh) = f_i(a+h) - f_i(a) - Df_i(a)h = \tilde{f}_i(h)$, so that

$$\lim_{h \to 0} \frac{\tilde{f}_i(h)}{\|h\|} = 0 \quad \forall\, i$$

$$\implies \lim_{h \to 0} \frac{f(a+h) - f(a) - Lh}{\|h\|} = 0$$

which proves that f is differentiable at a and has derivative $Df(a) = L$. $\square$

Note the following:

- Because of this theorem, it is enough to study differentiability of maps $f : \mathcal{O}_n \to \mathbb{R}$.
- Let $\gamma : \mathcal{O}_1 \to \mathbb{R}^n$. Then, γ is differentiable at $a \in \mathcal{O}_1$ iff γ_i is differentiable at a for all i and in that case,

$$D\gamma(a) = \gamma'(a) = \begin{pmatrix} \gamma_1'(a) \\ \gamma_2'(a) \\ \vdots \\ \gamma_n'(a) \end{pmatrix}$$

  This is the notion of velocity vector of a curve in $\mathbb{R}^n$ from elementary calculus.

5.2 Further properties of differentiable functions

Theorem 5.2.1 (Differentiability implies continuity)

Let $f : \mathcal{O}_n \to \mathbb{R}^m$. If f is differentiable at $a \in \mathcal{O}_n$, f is continuous at a.

Proof. Let f be differentiable at a. Then, for all $x \in \mathcal{O}_n$,

$$0 \leq \|f(x) - f(a)\| \leq \|f(x) - f(a) - Df(a)(x-a)\| + \|Df(a)(x-a)\|$$

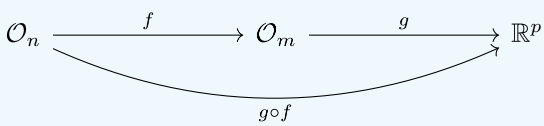
Taking the limit as $x \to a$, the first term on the right goes to 0 by the definition of $Df(a)$ and the second term goes to zero as any linear map is continuous. Hence, by the squeeze theorem,

$$\lim_{x \to a} \|f(x) - f(a)\| = 0$$

which proves that f is continuous at a. □

Theorem 5.2.2 (Chain rule)

Consider maps f, g such that



Assume that f is differentiable at $a \in \mathcal{O}_n$ and g is differentiable at $f(a) \in \mathcal{O}_m$. Then $g \circ f$ is differentiable at a and further

$$\underbrace{D(g \circ f)(a)}_{\mathbb{R}^n \to \mathbb{R}^p} = \underbrace{Dg(f(a))}_{\mathbb{R}^m \to \mathbb{R}^p} \cdot \underbrace{Df(a)}_{\mathbb{R}^n \to \mathbb{R}^m}$$

Proof. Let $A = Df(a)$ and $B = Dg(b)$ where $b = f(a)$. For x, y in sufficiently small neighbourhoods of a, b respectively, we consider the maps

$$\begin{aligned} r_f(x) &= f(x) - f(a) - A(x-a) \\ r_g(y) &= g(y) - g(b) - B(y-b) \\ r(x) &= g(f(x)) - g(b) - BA(x-a) \end{aligned}$$

By definition of the derivative,

$$\lim_{x \to a} \frac{r_f(x)}{\|x-a\|} = 0 \qquad \lim_{y \to b} \frac{r_g(y)}{\|y-b\|} = 0$$

We wish to prove that

$$\lim_{x \to a} \frac{r(x)}{\|x-a\|} = 0$$

We have,

$$\begin{aligned} r(x) &= g(f(x)) - g(b) - BA(x-a) \\ &= g(f(x)) - g(b) + B(r_f(x) - f(x) + f(a)) \\ &= Br_f(x) + g(f(x)) - g(b) - B(f(x) - f(a)) \\ \implies r(x) &= Br_f(x) + r_g(f(x)) \end{aligned}$$

Now,

$$\lim_{x \to a} \frac{Br_f(x)}{\|x-a\|} = B\left(\lim_{x \to a} \frac{Br_f(x)}{\|x-a\|}\right) = 0$$

The other term requires some more analysis. Fix $\varepsilon > 0$. There is $\delta > 0$ such that $\|r_g(y)\| < \varepsilon\|y-b\|$ for all y with $0 < \|y-b\| < \delta$. By continuity of f at a, there is $\delta_1 > 0$ such that $\|f(x) - f(a)\| < \delta$ for all x with $0 < \|x-a\| < \delta_1$. Hence,

$$0 < \|x-a\| < \delta_1 \implies \|r_g(f(x))\| < \varepsilon\|f(x) - f(a)\|$$

and so

$$\lim_{x \to a} \frac{r_g(f(x))}{\|f(x) - f(a)\|} = 0$$

We now note,

$$\|f(x) - f(a)\| = \|r_f(x) + A(x-a)\| \leq \|r_f(x)\| + M_A\|x-a\|$$

where we get $M_A > 0$ by Theorem **??**. Hence, we finally have

$$\lim_{x \to a} \frac{r(x)}{\|x-a\|} = \lim_{x \to a} \frac{r_g(f(x))}{\|f(x) - f(a)\|} \frac{\|f(x) - f(a)\|}{\|x-a\|} = 0$$

which completes the proof. □