Contents

Chapter I

The multivariate derivative

Convention. *Throughout the document, V, W will stand for generic normed linear spaces, and Ω, Υ will denote open subsets of V, W respectively. $\mathscr{D}$ will denote a domain of V.*

Convention. *A result that involves $\mathbb{K}$ (or that doesn't mention the field to be $\mathbb{R}$ or $\mathbb{C}$ explicitly) will actually stand for two results, one for $\mathbb{K} := \mathbb{R}$, and one for $\mathbb{K} := \mathbb{C}$.*

1 The Fréchet derivative

January 10, 2023

Remark. *Normed linear spaces will be over $\mathbb{K}$, and we'll also view them as metric and topological spaces.*

Definition 1.1 (Fréchet derivative)**.** Let V, W be normed linear spaces and $f \colon \Omega \to W$ with Ω being open in V.[1] Then we say that a linear map[2] $L \colon V \to W$ is a (Fréchet) derivative of f at a point $c \in \Omega$ iff for every $\varepsilon > 0$, there exists a $\delta > 0$ such that

$$\|f(x) - f(c) - L(x - c)\| < \varepsilon \|x - c\|$$

[1] Note that for a nonzero normed linear space, $\Omega \subsetneq \ell(\Omega)$ as singletons are not open.
[2] Usual definitions require L to be bounded. See Proposition 1.7.

whenever $x \in B_\delta(c) \cap \Omega \setminus \{c\}$.

If such an L exists, we say that f is (Fréchet) differentiable at c.

If $f\colon \Omega \to S$ where $S \subseteq W$, then we say that f is (Fréchet) differentiable at c iff $\iota_{W \leftarrow S} \circ f$ is differentiable at c.

Remark. *Note that V, W have to be over a common field: The Fréchet derivative is a* linear map.

Hence, whenever a Fréchet derivative (or differentiability) will be mentioned in the hypotheses of a theorem, implicit will be the assumption that the spaces are over a common field.

Corollary 1.2. *The Fréchet derivative of a linear map is itself.*

Remark. *Hence, Fréchet differentiability need not imply continuity for there are unbounded linear operators (only in infinite dimensions though): Consider $V := \bigoplus_{i=1}^{\infty} \mathbb{R}$ over $\mathbb{R}$ with* sup *norm, and then consider the map $T\colon V \to \mathbb{R}$ given by*

$$e_n \mapsto n.$$

Lemma 1.3. *Let V, W be normed linear spaces and $T\colon V \to W$ be linear such that for every $\varepsilon > 0$, there exists a $\delta > 0$ such that*

$$\|Tx\| < \varepsilon \|x\|$$

for all $x \in B_\delta(0) \setminus \{0\}$. Then $T = 0$.

Proposition 1.4. *There exists at most one Fréchet derivative at any point.*

Notation. *This allows to denote it by $Df(c)$. In case of $f\colon \Omega \to S$ where $S \subseteq W$, we'll again use $Df(c)$, but to denote $D(\iota_{W \leftarrow S} \circ f)(c)$.*

Remark. *Unless stated otherwise, the norm considered on $\mathbb{K}^n$ will the l_2 norm (which can be induced from the Euclidean inner-product).*

However, it doesn't matter if we are only interested in Fréchet differentiability as norms in finite dimensions are equivalent so that we can use Proposition *1.6.*

Proposition 1.5 (Equivalence with differentiability on $\mathbb{R}$)**.** *Let Ω be open in $\mathbb{R}$ and $f\colon \Omega \to \mathbb{R}$ with $c \in \Omega$. Then f is differentiable at $c \iff f$ is Fréchet differentiable at c, in which case, we have*

$$Df(c)\colon x \mapsto f'(c)x.$$

Proposition 1.6 (Equivalent norms preserve derivative)**.** *Let V, W be vector spaces, each equipped with a pair of equivalent norms, unprimed and primed. Let $f\colon \Omega \to W$ where Ω is open in W. Then f is Fréchet differentiable at $c \in \Omega$ in the unprimed norms $\iff$ it is so in the primed norms, in which case,*

$$Df(c) = D'f(c).$$

Remark. *Since norms in finite dimensions are equivalent, this means that the Fréchet derivative is independent of the norm for finite-dimensional spaces!*

Proposition 1.7. *A function with a bounded Fréchet derivative at a point is continuous at that point.*

Remark. *Since linear maps from finite-dimensional domains are bounded, we don't need to worry about boundedness of the Fréchet derivative in finite-dimensional domains.*

2 Directional derivatives

January 10, 2023

Definition 2.1 (Directional derivatives)**.** Let V, W be normed linear spaces and $f\colon \Omega \to W$ where Ω is open in V. Let $c \in \Omega$ and $v \in V \setminus \{0\}$. Take an $\varepsilon > 0$ such that $c + tv \in \Omega$ for each $t \in (-\varepsilon, \varepsilon)$. Then, if existent, we define[3]

$$D_v f(c) := \lim_{t \to 0} \frac{f(c+tv) - f(c)}{t\|v\|}.$$

If this exists, then we say that f is differentiable along v and call $D_v f(c)$ its derivative.

Again we define things for $f\colon \Omega \to S$ for $S \subseteq W$, as before.

[3] On the right-hand-side, the function is on $(-\varepsilon, \varepsilon) \setminus \{0\} \to W$. Note that 0 is a limit point of $(-\varepsilon, \varepsilon) \setminus \{0\}$.

Remark. *The above limit won't depend on ε, and hence we don't mention it on the left-hand-side.*

Also, here, unlike for the Fréchet derivative, V and W can be over differ-ent *fields!*

Corollary 2.2 (Directional derivatives are blind to magnitude)**.** *Let V, W be normed linear spaces and $f\colon \Omega \to W$ where Ω is open in V. Let f be differentiable at a point $c \in \Omega$ along a nonzero v. Then f is also differentiable at c along λv for any nonzero scalar λ with*

$$D_v f(c) = D_{\lambda v} f(c).$$

Proposition 2.3 (Fréchet differentiable $\implies$ differentiable along all directions)**.** *Let V, W be normed linear spaces and $f\colon \Omega \to W$, with Ω open in V, be Fréchet differentiable at $c \in \Omega$. Then f is differentiable along any $v \in V \setminus \{0\}$, and*

$$D_v f(c) = \frac{(Df(c))(v)}{\|v\|}.$$

Example 2.4 (The function needn't be differentiable even if all the directional derivatives exist!)**.** Consider $f\colon \mathbb{R}^2 \to \mathbb{R}$ given by

$$f(x,y) := \begin{cases} \frac{x^2 y}{x^4+y^2}, & (x,y) \neq (0,0) \\ 0, & (x,y) = (0,0) \end{cases}.$$

Then f is differentiable along all directions at $(0,0)$ but not even continuous at $(0,0)$, let alone differentiable![4]

Example 2.5 (The function needn't be differentiable even if the derivatives along all the directions come from some L!)**.** Define $f\colon \mathbb{R}^2 \to \mathbb{R}$ by

$$(x,y) \mapsto \begin{cases} \frac{x^3 y}{x^4+y^2}, & (x,y) \neq (0,0) \\ 0, & (x,y) = (0,0) \end{cases}.$$

Then for all $v \in \mathbb{R}^2 \setminus \{(0,0)\}$, we have[5]

$$\begin{aligned} D_v f(0,0) &= 0 \\ &= 0v \end{aligned}$$

and still f is not differentiable at $(0,0)$, let alone 0 being its Fréchet derivative.

[4] Differentiability $\implies$ continuity in finite-dimensions.

[5] The 0 here is the zero map.

Proposition 2.6 (Scaling of directional derivatives under equivalent norms)**.** *Continuing Proposition 1.6, f is differentiable at c along a nonzero v in the unprimed norms $\iff$ it is so in the primed norms, in which case,*

$$\|v\| D_v f(c) = \|v\|' D'_v f(c).$$

3 Partial derivatives

January 12, 2023

Definition 3.1 (Partial derivatives)**.** Let V, W be normed linear spaces with bases $\mathcal{B}$ and $\mathcal{C}$ respectively. Let $f\colon \Omega \to W$ where Ω is open in V. Let $\tilde{e} \in \mathcal{C}$ and $e \in \mathcal{B}$. Let $f_{\tilde{e}}\colon \Omega \to \mathbb{K}$ be the $\tilde{e}$-component of f. Then for $c \in \Omega$, if existent, we define

$$\begin{aligned} \partial_e f_{\tilde{e}}(c) &:= \lim_{t\to 0} \frac{f_{\tilde{e}}(c+te) - f_{\tilde{e}}(c)}{t} \\ &= \|e\| D_e f_{\tilde{e}}(c) \\ &\overset{(*)}{=} D f_{\tilde{e}}(c)(e) \end{aligned}$$

where the starred equality holds[6] if $f_{\tilde{e}}\colon \Omega \to \mathbb{K}$ is Fréchet differentiable.

Remark. *The basis of W should technically have been incorporated in the notation.*

Partials also can be defined even if V, W have different bases!

Theorem 3.2 (Jacobian)**.** *Let V, W be finite-dimensional normed linear spaces with ordered bases $\mathcal{B}$ and $\mathcal{C}$ respectively. Let Ω be open in V and $f\colon \Omega \to W$ be differentiable at $c \in \Omega$. Then all the partials exist at c, and we have*[7]

$$\big[Df(c)\big]_{\mathcal{C}\leftarrow\mathcal{B}} = \begin{bmatrix} \partial_1 f_1(c) & \cdots & \partial_n f_1(c) \\ \vdots & & \vdots \\ \partial_1 f_m(c) & \cdots & \partial_n f_m(c) \end{bmatrix}.$$

[6] Makes sense even!

[7] f_i's are as in Definition 3.1.

4 Some results

January 20, 2023

Theorem 4.1 (Differentiability in terms of components). *Let V, W be normed linear spaces with W being finite-dimensional with a basis $\mathcal{C}$. Let $f\colon \Omega \to W$, where Ω is an open subset of V, decompose into $f_{\tilde{e}}\colon \Omega \to \mathbb{K}$ for $\tilde{e} \in \mathcal{C}$. Then f is differentiable at a point $c \in \Omega \iff$ each $f_{\tilde{e}}$ is differentiable at c, in which case,*

$$Df(c)(v) = \sum_{\tilde{e}\in\mathcal{C}} Df_{\tilde{e}}(c)(v)\,\tilde{e}.$$

Theorem 4.2 (Chain rule). *Let U, V, W be normed linear spaces. Let Ω be open in U and Υ be open in V. Let $f\colon \Omega \to \Upsilon$ be differentiable at $c \in \Omega$ and $g\colon \Upsilon \to W$ be differentiable at $f(c) \in \Upsilon$. Then $g \circ f$ is differentiable at c with*

$$D(g \circ f)(c) = Dg(f(c)) \circ Df(c).$$

Corollary 4.3. *Let V, W be finite-dimensional normed linear spaces and $f\colon \Omega \to \Upsilon$ be invertible where Ω, Υ are open in V, W respectively. Let f be differentiable at $c \in \Omega$ and f^{-1} be differentiable at $f(c) \in \Upsilon$. Then the following hold:*

(i) $\dim V = \dim W$.

(ii) $Df(c)\colon V \to W$ *is invertible with* $(Df(c))^{-1} = Df^{-1}(f(c))$.

Theorem 4.4 (Continuous partials $\implies$ differentiability). *Let V, W be finite-dimensional normed linear spaces over $\mathbb{R}$. Let $f\colon \Omega \to W$ where Ω is open in V and let there exist a basis of V in which f is directionally differentiable throughout Ω along all directions, with the directional derivatives being continuous at a point $c \in \Omega$.*[8] *Then f is differentiable at c.*

Theorem 4.5 (Mean value for $\Omega \to \mathbb{R}$). *Let V, W be normed linear spaces with W being one-dimensional over $\mathbb{R}$. Let $f\colon \Omega \to W$ (Ω open) be continuous over $[x; y] \subseteq \Omega$ with $x \neq y$. Let $(\tilde{e})$ be a basis of W and $\partial_{y-x} f_{\tilde{e}}$ exist throughout $(x; y)$.* [9]*. Then there exists a point $\xi \in (x; y)$ such that*

$$f(y) - f(x) = \partial_{y-x} f_{\tilde{e}}(\xi).$$[10]

[8] This is equivalent to demanding the continuity of all the partials at c.

[9] The continuity on $(x; y)$ doesn't follow from differentiability on $(x; y)$ for V might be infinite dimensional.

[10] If f is differentiable on $(x; y)$, then this can be written as $Df_{\tilde{e}}(\xi)(y - x)$.

Corollary 4.6 (Zero derivative $\implies$ constant)**.** *Let V, W be normed linear spaces over $\mathbb{R}$ with W being one-dimensional. Let $f\colon \Omega \to W$, where Ω is open and connected in V, be differentiable throughout with*[11] *$Df(x) = 0$ for each $x \in \Omega$. Then f is constant over Ω.*

Theorem 4.7 (Mean value for $\Omega \to \mathbb{R}^n$)**.** *Let V, W be normed linear spaces over $\mathbb{R}$ with W being finite-dimensional. Let $f\colon \Omega \to W$, where Ω is open in V, be differentiable on $(x; y)$ and continuous on $[x; y] \subseteq \Omega$. Let $u \in W$ and $\mathcal{C}$ be a basis of W. Then there exists a point $\xi \in (x; y)$ such that*

$$\left\langle [u]_{\mathcal{C}}, [f(y) - f(x)]_{\mathcal{C}} \right\rangle = \left\langle [u]_{\mathcal{C}}, [Df(\xi)(y - x)]_{\mathcal{C}} \right\rangle$$

where $\langle \cdot, \cdot \rangle$ is the Euclidean inner-product on $\mathbb{R}^{|\mathcal{C}|}$.

Further, if V is finite-dimensional too with basis $\mathcal{B}$, then

$$\left\| [f(y) - f(x)]_{\mathcal{C}} \right\|_2 \leq M \left\| [y - x]_{\mathcal{B}} \right\|_2$$

where $\|\cdot\|_2$ represents the Euclidean norm on the respective spaces, and[12]

$$M := \sup_{z \in \Omega} \left\| [Df(z)]_{\mathcal{C} \leftarrow \mathcal{B}} \right\|$$

is existent.

5 Higher partials

January 25, 2023

Definition 5.1 (Higher partials)**.** Let V, W be normed linear spaces and $f\colon \Omega \to W$ where Ω is open in V. Let $\mathcal{B}$ and $\mathcal{C}$ be bases of V and W respectively. Then we define higher partials inductively as follows:

Let $\tilde{e} \in \mathcal{C}$ and $f_{\tilde{e}}\colon \Omega \to \mathbb{K}$ be f's $\tilde{e}$-component with respect to basis $\mathcal{C}$. Firstly, we define

$$\partial_{()} f_{\tilde{e}} := f_{\tilde{e}}.$$

Now, suppose $\partial_{e_1, \ldots, e_n} f_{\tilde{e}}\colon \mathscr{U} \to \mathbb{K}$ has been defined for $e_1, \ldots, e_n \in \mathcal{B}$ where $\mathscr{U}$ is open in V. Then, for another $e \in \mathcal{B}$, we define

$$\partial_{e, e_1, \ldots, e_n} f_{\tilde{e}}\colon \mathscr{V} \to \mathbb{K}$$

[11] 0 here is the zero function $V \to W$.

[12] The norms of the matrices are taken with $\mathbb{R}^k$'s under the Euclidean norm.

where $\mathscr{V}$ is the interior of the the set of all the points c in $\mathscr{U}$ wherever $\partial_e(\partial_{e_1,\dots,e_n} f_{\tilde{e}})_{(1)}(c)$ exists,[13] and we define

$$\partial_{e,e_1,\dots,e_n} f_{\tilde{e}}(c) := \partial_e(\partial_{e_1,\dots,e_n} f_{\tilde{e}})_{(1)}(c)$$

for $c \in \mathscr{V}$.

Remark. *This leads to a notational collision for the case of $\partial_e f_{\tilde{e}}$ which we have defined before (in Definition 3.1). But it's mild.*

Definition 5.2 (Differentiability classes (with respect to bases))**.** Let V, W be normed linear spaces and $f : \Omega \to W$ where Ω is open in V. Let $\mathcal{B}$ and $\mathcal{C}$ be bases of V and W respectively. Then we say that f is of differentiability class C^n for $n \geq 0$ iff for any $e_1, \dots, e_n \in \mathcal{B}$ and any $\tilde{e} \in \mathcal{C}$, the following hold:

(i) The domain of $\partial_{e_1,\dots,e_n} f_{\tilde{e}}$ is whole of Ω.

(ii) $\partial_{e_1,\dots,e_n} f_{\tilde{e}}$ is continuous throughout Ω.

Remark. *Note that the differentiability classes have to be talked of in the context of chosen bases!*

Lemma 5.3. *Let V, W be normed linear spaces and $f : \Omega \to W$ (Ω open) be C^n ($n \geq 0$) in the respective bases $\mathcal{B}$, $\mathcal{C}$. Let $e_1, \dots, e_k \in \mathcal{B}$ for $0 \leq k \leq n$. Then, for any $0 \leq l \leq k$ and any $\tilde{e} \in \mathcal{C}$, the following hold:*

(i) $\partial_{k,\dots,1} f_{\tilde{e}}$ is C^{n-k} in bases $\mathcal{B}$ and (1).

(ii) $\partial_{k,\dots,l+1}(\partial_{l,\dots,1} f_{\tilde{e}})_{(1)} = \partial_{k,\dots,1} f_{\tilde{e}}$.

Theorem 5.4 (Clairaut)**.** *Let V be a normed linear space with a basis $\mathcal{B}$. Let $f : \Omega \to \mathbb{R}$ where Ω is open in V. Let $e_1, e_2 \in V$ be distinct. Let the domains of*[14] *$\partial_1 f_{(1)}$, $\partial_2 f_{(1)}$, $\partial_{1,2} f_{(1)}$ and $\partial_{2,1} f_{(1)}$ be whole of Ω with $\partial_{1,2} f_{(1)}$ and $\partial_{2,1} f_{(1)}$ being continuous at $c \in \Omega$. Then*

$$\partial_{1,2} f_{(1)}(c) = \partial_{2,1} f_{(1)}(c).$$

[13] The subscript (1) denotes the fact that we are taking the usual basis for the normed linear space $\mathbb{K}$, namely the singleton containing 1.

[14] See Footnote 13.

Proposition 5.5 (Generalized Clairaut). *Let V be a normed linear space with W being over $\mathbb{R}$ and $f\colon \Omega \to \mathbb{R}$ (Ω open) be C^n ($n \geq 0$) in the respective bases $\mathcal{B}$,* (1). *Then for any $e_1, \dots, e_k \in \mathcal{B}$ for $0 \leq k \leq n$, we have*

$$\partial_{1,\dots,k} f_{(1)} = \partial_{\sigma(1),\dots,\sigma(k)} f_{(1)}$$

for any permutation $\sigma \in S_k$.

Theorem 5.6 (Taylor). *Let V be a normed linear space over $\mathbb{R}$ and $f\colon \Omega \to \mathbb{R}$, with Ω open in V, be C^{n+1} ($n \geq 0$) in bases $(e_1, \dots, e_m)$ and* (1). *Let $[x; y] \subseteq \Omega$ with $x \neq y$. Then there exists a $\xi \in (x; y)$ such that*

$$f(y) = \sum_{|\alpha| \leq n} \frac{1}{\alpha!} \partial^\alpha f_{(1)}(x)(y-x)^\alpha + \sum_{|\alpha|=n+1} \frac{1}{\alpha!} \partial^\alpha f_{(1)}(\xi)(y-x)^\alpha$$

where α ranges over $\mathbb{N}^m$.[15]

[15] The usual multi-index notation is used. See Appendix 1.

Appendix A

1 Multi-index notation

January 27, 2023

Definition 1.1 (Multi-index factorials and partials)**.** Let $n \geq 0$ and $\alpha \in \mathbb{N}^n$. Then we define

$$\alpha! := \alpha_1! \cdots \alpha_n!, \text{ and}$$
$$|\alpha| := \alpha_1 + \cdots + \alpha_n.$$

For $x \in X^n$ where X is a set where a product (with identity) is defined, we define

$$x^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n}.$$

Also, for a multivariate $\mathbb{K}$-valued function f with the domain being finite-dimensional with a basis $(e_1, \ldots, e_n)$, we define

$$\partial^\alpha f_{(1)} := \partial_{\underbrace{1,\ldots,1}_{\alpha_1 \text{ times}},\ldots,\underbrace{n,\ldots,n}_{\alpha_n \text{ times}}} f_{(1)}.$$

Lemma 1.2 (Multinomial expansion)**.** *Let R be a commutative ring with identity and $x \in R^n$, $n \geq 0$. Then for $k \geq 0$, we have*

$$(x_1 + \cdots + x_n)^k = \sum_{|\alpha|=k} \frac{k!}{\alpha!} x^\alpha$$

where $\alpha \in \mathbb{N}^n$.