Chapter 4

Calculus in $\mathbb{R}^n$

4.1 Linear maps

Definition 4.1. *Suppose $\mathbb{V}$ and $\mathbb{W}$ are real vector spaces.*

1. *A function $f\colon \mathbb{V} \to \mathbb{W}$ is called a* ***linear mapping*** *if*

$$f(\alpha v_1 + \beta v_2) = \alpha f(v_1) + \beta f(v_2)$$

   for all $\alpha, \beta \in \mathbb{R}$ and $v_1, v_2 \in \mathbb{V}$.

2. *A linear mapping $f\colon \mathbb{V} \to \mathbb{W}$ is called* ***bounded*** *if there is some constant $C > 0$ such that $\|f(v)\| \leq C$ for all $v \in \mathbb{V}$ such that $\|v\| = 1$. The set of all bounded linear mappings $\mathbb{V} \to \mathbb{W}$ is denoted by $B(\mathbb{V}; \mathbb{W})$.*

3. *Suppose $f\colon \mathbb{V} \to \mathbb{W}$ is a bounded linear mapping. The* ***operator norm*** *of f is defined by*

$$\|f\| = \sup\{\|f(v)\| \mid \|v\| = 1\}.$$

Example 4.2.

1. *All linear functions $\mathbb{R}^n \to \mathbb{R}^m$ are bounded.*

2. *Let $U \subset \mathbb{R}$ be open. The function $f \mapsto f'$ is a bounded linear mapping $C^1(U) \to C^0(U)$. What is the operator norm of this mapping?*

3. *Let $w\colon [a, b] \to \mathbb{R}$ be a continuous function. The function*

$$u \mapsto \int_a^b u(t)\, w(t)\, dt$$

is a bounded linear mapping $C^0([a,b]) \to \mathbb{R}$. *What is the operator norm of this mapping?*

Remark 4.3. *All linear mappings in* $B(\mathbb{R}^n, \mathbb{R}^m)$ *take the form* $f(x) = Mx$, *where* M *is a* $m \times n$ *matrix. Here we make use of the standard bases of the Cartesian spaces. In these notes we slightly abuse notation, writing* $M \in B(\mathbb{R}^n, \mathbb{R}^m)$.

Proposition 4.4. *Suppose* $\mathbb{V}$ *and* $\mathbb{W}$ *are real vector spaces and that* $f\colon \mathbb{V} \to \mathbb{W}$ *is linear. Then* f *is continuous if and only if* f *is bounded.*

Proof. Suppose $f\colon \mathbb{V} \to \mathbb{W}$ is continuous, and in particular continuous at $0 \in \mathbb{V}$. Since $f(0) = 0$, there exists $\delta > 0$ such that if $\|v\| < \delta$ then $\|f(v)\| < 1$. Let $v \in \mathbb{V}$ be such that $\|v\| = 1$. Then $\|\frac{1}{2\delta}v\| < \delta$ and thus

$$\frac{1}{2\delta}\|f(v)\| = \left\| f\left(\frac{1}{2\delta}v\right)\right\| < 1.$$

By linearity, this is equivalent to $\|f(v)\| < 2\delta$. Thus we see that f is bounded.

Suppose now that f is bounded, with $\|f(v)\| \leq C$ for all $v \in \mathbb{V}$ with $\|v\| = 1$. For any distinct $v_1, v_2 \in \mathbb{V}$ we have

$$\begin{aligned} \|f(v_1) - f(v_2)\| &= \|f(v_1 - v_2)\| \\ &= \|v_1 - v_2\| \left\| f\left(\frac{v_1 - v_2}{\|v_2 - v_2\|}\right)\right\| \leq C\|v_1 - v_2\|. \end{aligned}$$

Thus f is Lipschitz continuous. □

Proposition 4.5. *Suppose* $\mathbb{V}$ *and* $\mathbb{W}$ *are Banach spaces. Then* $B(\mathbb{V}, \mathbb{W})$ *is a Banach space relative to the operator norm.*

Proof. It is straightfoward to see that $B(\mathbb{V}, \mathbb{W})$ is a vector space. To see that it is complete, let $\{f_n\}_{n=1}^{\infty}$ be a Cauchy sequence in $B(\mathbb{V}, \mathbb{W})$.

Let $\varepsilon > 0$ and fix $v \in \mathbb{V} \setminus \{0\}$. Then there exists $N \in \mathbb{N}$ such that when $n, m \geq N$ we have $\|f_n - f_m\| < \varepsilon/\|v\|$. Thus

$$\|f_n(v) - f_m(v)\| = \|v\| \left\| (f_n - f_m)\left(\frac{v}{\|v\|}\right)\right\| \leq \|v\|\|f_n - f_m\| < \varepsilon.$$

Thus $\{f_n(v)\}_{n=1}^{\infty}$ is a Cauchy sequence in $\mathbb{W}$, and therefore converges to some $w \in \mathbb{W}$. We define $f\colon V \to W$ by $f(v) = w$.

We claim that the map f is linear. To see this, let $v_1, v_2 \in \mathbb{V}$ and $\alpha_1, \alpha_2 \in \mathbb{R}$. For any $n \in \mathbb{N}$ we have

$$\begin{aligned}\|f(\alpha_1 v_1 + \alpha_2 v_2) - \alpha_1 f(v_1) - \alpha_2 f(v_2)\| \\ \leq \|f(\alpha_1 v_1 + \alpha_2 v_2) - f_n(\alpha_1 v_1 + \alpha_2 v_2)\| \\ + |\alpha_1|\|f_n(v_1) - f(v_1)\| + |\alpha_2|\|f_n(v_2) - f(v_2)\|.\end{aligned}$$

Since the limit of the right side is zero, we see that

$$f(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 f(v_1) + \alpha_2 f(v_2).$$

To see that $f_n \to f$ with respect to the operator norm, fix $\varepsilon > 0$ and let $N \in \mathbb{N}$ such that $n, m \geq N$ imply $\|f_n - f_m\| < \varepsilon/2$. We claim that if $n \geq N$ then $\|f - f_n\| < \varepsilon$.

To see this, consider $v \in \mathbb{V}$. For every $m \geq N$ we have

$$\|f(v) - f_n(v)\| \leq \|f(v) - f_m(v)\| + \|f_m(v) - f_n(v)\| \leq \|f(v) - f_m(v)\| + \frac{\varepsilon}{2}.$$

Since $\|f(v) - f_m(v)\| \to 0$ as $m \to \infty$, it must be that

$$\|f(v) - f_n(v)\| \leq \frac{\varepsilon}{2}.$$

Since this last estimate holds for each v, and since n was chosen independent of v, we have established that

$$\|f - f_n\| < \varepsilon$$

for all $n \geq N$. Thus $f_n \to f$ with respect to the operator norm.

Finally, we now that convergence with respect to the operator norm implies that f is bounded. Thus $B(\mathbb{V}, \mathbb{W})$ is indeed a Banach space. □

Example 4.6. *Let $U \subset \mathbb{R}^n$. Then $C^0(U; B(\mathbb{R}^n; \mathbb{R}^m))$ is a Banach space.*

4.2 Differentiation

Definition 4.7. *Let $U \subset \mathbb{R}^n$ be open. A function $f \colon U \to \mathbb{R}^m$ is* ***differentiable at $x \in U$*** *if there exists $M \in B(\mathbb{R}^n; \mathbb{R}^m)$ such that*

$$\lim_{y \to 0} \frac{\|f(x+y) - f(x) - My\|}{\|y\|} = 0.$$

If f is differentiable at each $x \in U$ then we simply say that ***$f \colon U \to \mathbb{R}^m$ is differentiable****.*

Lemma 4.8. *Suppose $U \subset \mathbb{R}^n$ is open and $f : U \to \mathbb{R}^m$ is differentiable at each $x \in U$. Then there exists only one linear function $L : \mathbb{R}^n \to \mathbb{R}^m$ such that*

$$\lim_{y \to 0} \frac{\|f(x+y) - f(x) - My\|}{\|y\|} = 0.$$

Thus $x \to M$ is a function $U \mapsto B(\mathbb{R}^n; \mathbb{R}^m)$, which we denote by Df.

Proof. The proof is identical to the proof in the one-dimensional setting. □

Definition 4.9. *Suppose $U \subset \mathbb{R}^n$ is open and $f : U \to \mathbb{R}^m$ is differentiable. We say that f is **continuously differentiable**, and write $f \in C^1(U)$, if the function $x \mapsto Df(x)$ is a continuous map $U \to B(\mathbb{R}^n; \mathbb{R}^m)$.*

Lemma 4.10. *Suppose $U \subset \mathbb{R}^n$ is open and $f, g : U \to \mathbb{R}^m$ are differentiable at $x \in U$. Then*

1. *f is continuous at x;*
2. *$f + g$ is differentiable at x, with $D(f+g)(x) = Df(x) + Dg(x)$; and*
3. *for each $a \in \mathbb{R}$ the function af is differentiable at x with $D(af)(x) = aDf(x)$.*

Proof. Suppose $x_k \to x$ in $\mathbb{R}^n$; we write $x_k = x + y_k$ such that $y_k \to 0$. By f being differentiable, we see that

$$\|f(x + y_k) - f(x) - My_k\| \to 0.$$

This easily implies that

$$\|f(x_k) - f(x)\| \to 0,$$

and thus that f is continuous at x.

The linearity of $f \mapsto Df$ follows immediately from the definitions involved. □

Proposition 4.11 (The chain rule)**.** *Suppose $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^m$ are open. Suppose also that $g : U \to V$ is differentiable at x_*, and that $f : V \to \mathbb{R}^l$ is differentiable at $y_* = g(x_*)$. Then $f \circ g : U \to \mathbb{R}^l$ is differentiable at x_* with $D(f \circ g)(x_*) = Df(y_*) \circ Dg(x_*)$.*

Proof. Let $\varepsilon > 0$. Let $\mu > 0$ be such that whenever $y \in \mathbb{R}^l$ satisfies $\|y\| < \mu$ then

$$\|f(y_* + y) - f(y_*) - Df(y_*)y\| \leq \frac{\varepsilon}{2(1 + \|Dg(x_*)\|)}\|y\|.$$

Let $\delta_f > 0$ be such that if $x \in \mathbb{R}^n$ satisfies $\|x\| < \delta_f$, then $x_* + x \in U$ and $\|g(x_* + x) - g(x_*)\| < \mu$. Thus if $\|x\| < \delta_f$ then

$$\|f(g(x_* + x)) - f(g(x_*) - Df(y_*)(g(x_* + x) - g(x_*))\| < \frac{\varepsilon}{2(1 + \|Dg(x_*)\|)}\|g(x_* + x) - g(x_*)\|$$

Using the differentiability of g, fix $\delta_1 > 0$ such that $\|x\| < \delta_1$ implies $x_* + x \in U$ and

$$\|g(x_* + x) - g(x_*) - Dg(x_*)x\| \leq \|x\|.$$

Thus if $\|x\| < \delta_1$ we have

$$\|g(x_* + x) - g(x_*)\| \leq (1 + \|Dg(x_*)\|)\|x\|.$$

Thus if $\|x\| < \min(\delta_1, \delta_f)$ we have

$$\|f(g(x_* + x)) - f(g(x_*) - Df(y_*)(g(x_* + x) - g(x_*))\| < \frac{\varepsilon}{2}\|x\|.$$

The differentiability of g also implies that there exists $\delta_g > 0$ such that if $\|x\| < \delta_g$ then $x_* + x \in U$ and

$$\|g(x_* + x) - g(x_*) - Dg(x_*)x\| \leq \frac{\varepsilon}{2\|Df(y_*)\|}\|x\|.$$

Thus if $\|x\| < \delta_g$ then

$$\|Df(y_*)(g(x_* + x) - g(x_*) - Dg(x_*)x)\| < \frac{\varepsilon}{2}\|x\|.$$

Suppose now that $x \in \mathbb{R}^n$ is such that $\|x\| < \min(\delta_f, \delta_g, \delta_1)$. Then from the triangle inequality and the estimates above we have

$$\begin{aligned}
&\|f(g(x_* + x)) - f(g(x_*) - Df(y_*)Dg(x_*)x\| \\
&\quad \leq \|f(g(x_* + x)) - f(g(x_*) - Df(y_*)(g(x_* + x) - g(x_*))\| \\
&\qquad + \|Df(y_*)(g(x_* + x) - g(x_*) - Dg(x_*)x)\| < \varepsilon\|x\|.
\end{aligned}$$

This implies the desired result. $\square$

Definition 4.12. *Let $U \subset \mathbb{R}^n$ be open and $x = (x^1, \ldots, x^n) \in U$. A function $f : U \to \mathbb{R}^m$ is called **partially differentiable at x** if for $1 \le j \le n$ the functions $y \mapsto f(x^1, \ldots, x^{j-1}, x^j + y, x^{j+1}, \ldots, x^n)$ are differentiable at $y = 0$. These derivatives are called the **partial derivatives** of f and are denoted $\partial_j f(x)$.*

If f is partially differentiable at each $x \in U$, then we say that f is partially differentiable.

Theorem 4.13 (Partial derivative theorem)**.** *Let $U \subset \mathbb{R}^n$ be open.*

1. *A function $f : U \to \mathbb{R}^m$ is continuously differentiable if and only if it is partially differentiable with all partial derivatives continuous.*
2. *If $f : U \to \mathbb{R}^m$ is continuously differentiable, then for each $x \in U$ and $v = (v^1, \ldots, v^n) \in \mathbb{R}^n$ we have*

$$Df(x)v = \partial_1 f(x)v^1 + \cdots + \partial_n f(x)v^n.$$

Proof. Suppose that $f : U \to \mathbb{R}^m$ is continuously differentiable. Fix $x = (x^1, \ldots, x^n) \in U$ and integer j with $1 \le j \le n$. For $y \in \mathbb{R}$, define $x(y) \in \mathbb{R}^n$ by

$$x(y) = (x^1, \ldots, x^{j-1}, x^j + y, x^{j+1}, \ldots, x^n).$$

If y is sufficiently small, then $x(y) \in U$. Thus we can apply the chain rule to $f(x(y))$, deducing that f is partially differentiable. That f is continuously differentiable implies that the partial derivatives are each continuous.

Suppose now that f is partially differentiable with continuous partial derivatives. Fix $x = (x^1, \ldots, x^n) \in U$ and choose $\delta_1 \in (0, 1)$ such that $B_\delta(x) \subset U$. Suppose $y = (y^1, \ldots, y^n) \in B_\delta(0)$, so that $x + y \in U$. We have

$$\begin{aligned} f(x+y) - f(x) = f(&x^1 + y^1, x^2, \ldots, x^n) \\ &- f(x^1, \ldots, x^n) \\ &+ f(x^1 + y^1, x^2 + y^2, x^3, \ldots, x^n) \\ &- f(x^1 + y^1, x^2, x^3, \ldots, x^n) \\ &\vdots \\ &+ f(x^1 + y^1, \ldots, x^{n-1} + y^{n-1}, x^n + y^n) \\ &- f(x^1 + y^1, \ldots, x^{n-1} + y^{n-1}, x^n). \end{aligned}$$

Let $\varepsilon > 0$. By f being partially differentiable, there exists $\delta_f \in (0, \delta_1)$ such that if $\|y\| < \delta_f$ then for each $1 \leq j \leq n$ we have

$$\begin{aligned}\|f(x^1 + y^1, \dots, x^{j-1} + y^{j-1}, x^j + y^j, x^{j+1}, \dots, x^n) \\ - f(x^1 + y^1, \dots, x^{j-1} + y^{j-1}, x^j, x^{j+1}, \dots, x^n) \\ - \partial_j f(x^1 + y^1, \dots, x^{j-1} + y^{j-1}, x^j, x^{j+1}, \dots, x^n) y^j\| \\ < \frac{\varepsilon}{4n}|y^j| \leq \frac{\varepsilon}{2n}\|y\|.\end{aligned}$$

Finally, the continuity of the partial derivatives implies that we may choose $\delta_D \in (0, \delta_f)$ such that for each $1 \leq j \leq n$ we have

$$\|\partial_j f(x^1 + y^1, \dots, x^{j-1} + y^{j-1}, x^j, x^{j+1}, \dots, x^n) - \partial_j f(x^1, \dots, x^n)\| < \frac{\varepsilon}{4n}.$$

Assembling these estimates, we find that

$$\|f(x + y) - f(x) - \partial_1 f(x) y^1 - \cdots - \partial_n f(x) y^n\| < \varepsilon\|y\|.$$

Thus f is differentiable with

$$Df(x)y = \partial_1 f(x) y^1 + \cdots + \partial_n f(x) y^n.$$

Since the partial derivatives are each continuous in x, we see that $x \mapsto Df(x)$ is continuous as well.

Finally, we note that in the course of establishing the first claim we have also established the second. □

Remark 4.14. *The definitions and results of this section remain valid if the Cartesian spaces are replaced by Banach spaces. See, for example, the discussion in the monograph* Postmodern Analysis *by Jürgen Jost.*