Vv285 Linear Algebra and Functions of Multiple Variables Review[1]

Shaoxiong Yao

University of Michigan

E-mail: yaosx@umich.edu

[1] *This work is based on lecture slides VV285 taught by Prof. Horst Hohberger at UM-SJTU Joint Institute in 2018 summer.*

Contents

PART

I

Convergence and Continuity

Definition 0.1. Let $(V, \|\cdot\|)$ be a normed vector space. Then

$$B_\varepsilon(a) := \{x \in V : \|x - a\| < \varepsilon\}, \quad a \in V, \varepsilon > 0, \tag{0.1}$$

is called an ***open ball*** of radius ε about a.

Notice that an open ball may not have an obvious shape like $V = \mathcal{P}_n$.

Definition 0.2. Let $(V, \|\cdot\|)$ be a normed vector space. A set $U \subset V$ is called ***open*** if for every $a \in U$ there exists an $\varepsilon > 0$ such that $B_\varepsilon(a) \subset U$.

Remark. We note that any point is contained in some open ball inside the open set.

Example.
(i) Any open ball $B_\varepsilon(a), \varepsilon > 0, a \in V$, is an open set.
(ii) The empty set $\emptyset \subset V$ is open. (This is true because ***vacuously true statement***.)
(iii) The entire space V is an open set in V.

We will see that open sets are fundamental to understand properties of continuous functions, convergence in vector space and much more.

Therefore, it becomes important to answer a basic question:

If a set is open in a vector space $(V, \|\cdot\|)$, is it also open if $\|\cdot\|$ is replaced by some other norm?

Definition 0.3. Let V be a vector space on which we may define two norms $\|\cdot\|_1$ and $\|\cdot\|_2$. Then the two noms are called ***equivalent*** is there exists two constants $C_1, C_2 > 0$ such that

$$C_1\|x\|_1 \le \|x\|_2 \le C_2\|x\|_1, \qquad \text{for all } x \in V. \tag{0.2}$$

Example. In $\mathbb{R}^n$ we have (amongst others) the following two possible choices of norms

$$\|x\|_2 := \left(\sum_{i=1}^{n} |x_i||^2\right)^{1/2}, \qquad \|x\|_\infty := \max_{1\le i\le n} |x_i|. \tag{0.3}$$

We note that

$$\begin{aligned}\|x\|_2 &\le \left(n\|x\|_\infty^2\right)^{1/2} = \sqrt{n}\|x_\infty\|,\\ \|x\|_\infty &\le \left(\|x\|_\infty^2 + \sum_{i\ne j, \|x_j\|=\|x\|_\infty} \|x_i\|\right)^{1/2} = \|x\|_2.\end{aligned}$$

then we have
$$\frac{1}{\sqrt{n}}\|x\|_2 \le \|x\|_\infty \le \|x\|_2.$$

Remark. We want to show that if two norms are equivalent, then a set is an open set with respect to both norms.
Assume U is an open set with respect to $\|\cdot\|_1$, then we have for $a \in U$, there exist $\varepsilon > 0$ such that $B_\varepsilon \subset U$. We then have
$$B_\varepsilon = \{x : \|x-a\|_1 < \varepsilon\}.$$
Then to claim that U is also an open set with respect to $\|\cdot\|_2$, we want to find $B_{\varepsilon'} \subseteq B_\varepsilon$ for all x above. We then find $\varepsilon' = C_1\varepsilon$ such that
$$\begin{aligned}\|x-a\|_2 \le \varepsilon' \Rightarrow C_1\|x-a\|_1 &\le \|x-a\|_2 \le C_1\varepsilon' \\ &\Rightarrow \|x-a\|_1 \le \varepsilon.\end{aligned}$$
We then have $B_{\varepsilon'} \subseteq B_\varepsilon$ and U is also an open set with respect to $\|\cdot\|_2$.

Definition 0.4. Let $(V, \|\cdot\|)$ be a normed vector space and (v_n) a sequence in V. Then (v_n) converges to a (unique) limit $v \in V$,
$$v_n \xrightarrow{n\to\infty} v \qquad \text{if and only if} \qquad \|v_n - v\| \xrightarrow{n\to\infty} 0.$$

For later use, we note:

Remark. If a sequence (v_n) in $(V, \|\cdot\|)$ converges to $v \in V$, then $\|v_n\| \to \|v\|$. This follows from
$$\big|\|v_n\| - \|v\|\big| \le \|v - v_n\| \to 0.$$

Remark. We notice that if two norms on a vector space are equivalent, then a sequence that converges to a limit with respect to the first norm is also convergent to the same limit with respect to the second norm.
For the first norm we have,
$$\forall\varepsilon > 0, \exists N \in \mathbb{N}, \|v_n - v\|_1 \le \varepsilon \text{ for all } n > N,$$
then for the second norm, $\forall\varepsilon > 0$, we can find $\frac{\varepsilon}{C_2} > 0$ such that
$$\exists N > 0, \|v_n - v\|_1 < \frac{\varepsilon_1}{C_2} \text{ for all } n > N.$$
We then have for this $N > 0$,
$$\|v_n - v\|_2 \le C_2\|v_n - v\|_1 < \varepsilon \text{ for all } n > N.$$

The following theorem is of fundamental importance:

Theorem 0.1. In a finite-dimensional vector space, all norms are equivalent.

A major consequence of **Theorem 0.1** is that if we have several norms at our disposal in a finite-dimensional spae, then we can freely choose a convenient one in order to show openness of sets, convergence of sequences, etc.

The proof of **Theorem 0.1** requires some preliminary work.

We recall two basic facts from the theory of sequences of real numbers:

- Every bounded and monotonic sequence of real numbers converges.

- Every sequence of real numbers has a monotonic subsequence.

Therefor, these yields the following fundamental result (cf. 186 **Theorem 2.2.35**):

Theorem 0.2. *(Theorem of Bolzano-Weierstraß)* Every bounded sequence of real numbers has a convergent subsequence.

We remark that the **Theorem of Bolzano-Weierstraß** easily implies that every Cauchy sequence of real numbers converges, because every Cauchy sequence that has a convergent subsequence must itself converge. Thus the basic ingredient in proving that the real numbers (with the usual metric) are complete is the fact that a bounded, monotonic sequences converges.

Theorem 0.3. *(Theorem of Bolzano-Weierstraß in $\mathbb{R}^n$)* Let $\left(x^{(m)}\right)_{m\in\mathbb{N}}$ be a sequence of vectors in $\mathbb{R}$, i.e., $x^{(m)} = \left(x_1^{(m)}, ..., x_n^{(m)}\right)$. Suppose that there exists a constant $C > 0$ such that $|x_k^{(m)}| < C$ for all $m \in \mathbb{N}$ and each $k = 1, ..., n$. Then there exists a subsequence $\left(x^{(m_j)}\right)_{j\in\mathbb{N}}$ that converges to a vector $y \in \mathbb{R}^n$ in the sense that

$$x_k^{(m_j)} \xrightarrow{j\to\infty} y_k \qquad \text{for } k = 1, ..., n.$$

Proof. Consider the real coordinate sequence $\left(x_1^{(m_j)}\right)_{m\in\mathbb{N}}$. By assumption, this sequence is bounded, so by the Theorem of Bolzano-Weierstraß (theorem on sequence of numbers.) there exists a convergent subsequence $(x_1^{m_{j_1}})$ with some limit, say $y_1 \in \mathbb{R}$.
The second coordinate sequence (x_2^m) is also bounded an have a convergent subsequence, but this subsequence does not need to have the same indices as that for $\left(x_1^{(m)}\right)$.
We therefore employ a trick: we consider the indices sequence m_{j_i} which gives a convergent sequence of x_1, then this will correspond to a subsequence of x_2 which is $\left(x_2^{(m_{j_1})}\right)$. For this sequence of numbers, we can find a subsequence $\left(x_2^{(m_{j_2})}\right)$ which converges to $y_2 \in \mathbb{R}$. Since $\left(x_1^{(m_{j_1})}\right)$ converges and m_{j_2} is a subsequence of m_{j_1}, we have $\left(x_1^{(m_{j_2})}\right)$ also converges to $y_1 \in \mathbb{R}$.
Similarly, we can find a sub-sub-subsequence with indices m_{j_3} and the third coordinate $\left(x_3^{(m_{j_3})}\right)$ will converge to some $y_3 \in \mathbb{R}$. We have the corresponding sub-sub-subsequence $(x_1^{m_{j_3}})$ and $(x_2^{m_{j_3}})$ will still converge to y_1 and y_2.
Repeating this procedure n times, the n-fold subsequence $(x_k^{m_{j_n}})$ converges to some $y_k \in \mathbb{R}$, $k = 1, ..., n$. Hence, the subsequence $(x^{(m_{j_n})})$ converges to some $y \in \mathbb{R}^n$. □

Lemma 0.4. Let $(V, \|\cdot\|)$ be a finite- or infinite-dimensional normed vector space and $\{v_1, ..., v_n\}$ an independent set in V. Then there exists a $C > 0$ such that for any $\lambda_1, ..., \lambda_n \in \mathbb{F}$

$$\left\|\lambda_1 v_1 + \cdots + \lambda_n v_n\right\| \geq C(|\lambda_1| + \cdots + |\lambda_n|). \tag{0.4}$$

Proof. Let $s := |\lambda_1| + \cdots + |\lambda_n|$. If $s = 0$, then all $\lambda_k = 0$ and the inequality (0.4) holds trivially for any C, so we can assume $s > 0$. Dividing by s, (0.4) becomes

$$\left\|\mu_1 v_1 + \cdots + \mu_n v_n\right\| \geq C, \quad \sum_{k=1}^{n} |\mu_k| = 1, \tag{0.5}$$

with $\mu_k = \lambda_k / s$.
Hence, we need to show

$$\underset{C>0}{\exists} \quad \underset{\substack{\mu_1,\ldots,\mu_n \in \mathbb{F} \\ |\mu_1|+\cdots+|\mu_n|=1}}{\forall} \left\|\mu_1 v_1 + \cdots + \mu_n v_n\right\| \geq C.$$

Suppose that this is false, i.e.,

$$\underset{C>0}{\forall} \quad \underset{\substack{\mu_1,\ldots,\mu_n \in \mathbb{F} \\ |\mu_1|+\cdots+|\mu_n|=1}}{\exists} \left\|\mu_1 v_1 + \cdots + \mu_n v_n\right\| < C.$$

In particular, choosing $C = 1/m$, $m = 1, 2, 3, ...$, we can find a sequence of vectors

$$u^{(m)} := \mu_1^{(m)} v_1 + \cdots + \mu_n^{(m)} v_n$$

such that $\|u^{(m)}\| \to 0$ as $m \to \infty$ and $|\mu_1^{(m)}| + \cdots + |\mu_n^{(m)}| = 1$ for all m.
Hence, for each $k = 1, ..., n$, $|\mu_k^{(m)}| \leq 1$ and so each coefficient sequence $\left(\mu_k^{(m)}\right)$ is bounded. Write

$$\mu^{(m)} := \left(\mu_1^{(m)}, ..., \mu_n^{(m)}\right)$$

By the Theorem of Bolzano Weierstraß in $\mathbb{R}^n$, there exists a subsequence of vectors $\left(\mu^{(m_j)}\right)_{j \in \mathbb{N}}$ that converges to some $\alpha = (\alpha_1, ..., \alpha_n) \in \mathbb{R}^n$. This corresponds to a subsequence $u^{(m_j)}$ of $u^{(m)}$ such that

$$u^{(m_j)} \xrightarrow{j \to \infty} \alpha_1 v_1 + \cdots + \alpha_n v_n =: u \qquad \text{with } |\alpha_1| + \cdots + |\alpha_n| = 1.$$

Since the vectors $v_1, ..., v_n$ are independent and not all α_k vanish, it follows that $u \neq 0$.
From previous remark, we note that if a sequence of vectors v_n converges to v, then its norm $\|v_n\|$ converges to $\|v\|$. We then have

$$\|u^{(m_j)}\| \xrightarrow{j \to \infty} \|u\| \neq 0.$$

But by our construction, $\left\|u^{(m)}\right\| \to 0$ as $m \to \infty$, so the subsequence $\left(\left\|u^{(m_j)}\right\|\right)$ must also converge to zero. This gives a contradiction. □

Remark. (1) The key point in this proof is to consider the contraposition which gives an inequality that can be interpreted as limit. We normalized the inequality first. We note that $\mu_1 v_1 + \cdots + \mu_n v_n$, $\sum_{k=1}^{n} |\mu_k|$ gives several hyperplanes in $\mathbb{R}^n$. These planes does not go through the origin and have a minimum distance to the origin.

(2) Here we need to note that $\{v_1, ..., v_n\}$ needs to a set of independent vectors. Otherwise we cannot guarantee that $\lambda_1 v_1 + \cdots + \lambda_n v_n \neq 0$.

(3) Also, here we use Theorem of Bolzano and Weierstraß in $\mathbb{R}^n$ which does not necessarily hold for infinite dimensional case.

We can now proceed to prove **Theorem 0.1**.

Proof. *(Theorem 0.1)* Let V be a finite-dimensional vector space, $\|\cdot\|$ be any norm on V and $\{v_1, ..., v_n\}$ a basis of V. Let $v \in V$ have the representation $v = \lambda_1 v_1 + \cdots + \lambda_n v_n$ with $\lambda_1, ..., \lambda_n \in \mathbb{F}$. By the triangle inequality,

$$\|v\| = \|\lambda_1 v_1 + \cdots + \lambda_n v_n\| \leq \sum_{i=1}^{n} |\lambda_i| \|v_1\| \leq C \sum_{i=1}^{n} |\lambda_i|$$

where $C := \max_{1 \leq i \leq n} \|v_i\|$ depends only on the basis and ont on v. We hence see that for any norm there are constants $C_1, C_2 > 0$ such that

$$C_1 \sum_{i=1}^{n} |\lambda_i| \leq \|v\| \leq C_2 \sum_{i=1}^{n} |\lambda|, \tag{0.6}$$

where the first inequality is just (0.4). Given two norms $\|\cdot\|_1$ and $\|\cdot\|_2$, it follows form there respective inequalities (0.6) that (0.2) holds. ☐

Remark. Note that if for $\|\cdot\|_1$, we have

$$C_{11} \sum_{i=1}^{n} |\lambda_i| \leq \|v\|_1 \leq C_{12} \sum_{i=1}^{n} |\lambda_i|,$$

where $C_{11}, C_{12} > 0$. For $\|\cdot\|_2$, we have

$$C_{21} \sum_{i=1}^{n} |\lambda_i| \leq \|v\|_2 \leq C_{22} \sum_{i=1}^{n} |\lambda_i|,$$

where $C_{21}, C_{22} > 0$. Then we have

$$\frac{C_{11}}{C_{22}} \|v\|_2 \leq \|v\|_1 \leq \frac{C_{12}}{C_{21}} \|v\|_2.$$

It is essential that **Theorem 0.1** assumes that V is a finite-dimensional vector space. In an infinite-dimensional space, it is possible to define non-equivalent norms.

Example. Consider the space of continuous functions on $[0, 1]$, $C([0, 1])$. We can define the two norms

$$\|f\|_\infty = \sup_{x \in [0,1]} |f(x)|, \quad \|f\|_1 = \int_0^1 |f(x)| dx.$$

We can show that the sequence of continuous functions $f_n = e^{-nx}$ converges with respect to $\|f\|_1$, but does not converge with respect to $\|f\|_\infty$.

We now introduce some important concepts.

Definition 0.5. Let $(V, \|\cdot\|)$ be a normed vector space and $M \subset V$.

(i) A point $x \in M$ is called an ***interior point of*** M is there exists an $\varepsilon > 0$ such that $B_\varepsilon(x) \subset M$.

(ii) The set of interior points of M is denoted by int M.

(iii) A point $x \in V$ is called a ***boundary point of*** M if for every $\varepsilon > 0$ $B_\varepsilon(x) \cap M \neq \emptyset$ and $B_\varepsilon(x) \cap (V \backslash M) \neq \emptyset$.

(iv) The set of boundary points of M is denoted by ∂M.

(v) A point that is neither a boundary nor an interior point of M is called an ***exterior point of*** M.

Remark. (i) An exterior point x of M is an interior point of $V\backslash M$. First, it is not an interior point, then for every $\varepsilon > 0$, $B_\varepsilon(x) \cap (V\backslash M) \neq \emptyset$. Then, it is not an boundary point, so there exists $\varepsilon > 0$, $B_\varepsilon(x) \cap M = \emptyset$ or $B_\varepsilon(x) \cap (V\backslash M) = \emptyset$. We note from the first argument, we can only have $B_\varepsilon(x) \cap M = \emptyset$. Then there exists $\varepsilon > 0$, $B_\varepsilon(x) \subset (V\backslash M)$. We have x is an interior point of $V\backslash M$.

(ii) For given M, any point of V is either an interior, boundary or exterior point of M.

Definition 0.6. Let $(V, \|\cdot\|)$ be a normed vector space and $M \subset V$. Then M is said to be ***closed*** if its complement $V\backslash M$ is open.

Remark. Of course, a set M does not need to be either open or closed. Some sets are open and closed at the same time.

Example. (i) A set consisting of a single point, $M = \{a\} \subset V$, is a closed set.

(ii) The empty set $\emptyset \subset V$ is closed. Also, from vacuously true, $\emptyset$ is open.

(iii) The entire space V is a closed set in V.

Lemma 0.5. Let $(V, \|\cdot\|)$ be a normed vector space and $M \subset V$.

(i) The set M is open if and only if $M = \text{int } M$.

(ii) The set M is closed if and only if $\partial M \subset M$.

Proof.

(i) We note that M is an open set iff all $x \in M$, there exists $\varepsilon > 0$ such that $B_\varepsilon(x) \subset M$. This is the definition that x is an interior point.

(ii) Suppose that M is closed. Then $V\backslash M$ is open. An open set can not contain a boundary point, since all its points are interior points. Hence, $\partial M \cap (V\backslash M) = \emptyset$ and so $\partial M \subset M$.

Suppose that $\partial M \subset M$. Then $V\backslash M$ contains only exterior points of M. But an exterior point of M is an interior point of $V\backslash M$, so $V\backslash M$ is open. Hence, M is closed.

□

Remark. We note that for any set $M \subseteq V$, int M, ∂M and int $(V\backslash M)$ is a partition of V.

Definition 0.7. Let $(V, \|\cdot\|)$ be a normed vector space and $M \subset V$. Then

$$\overline{M} := M \cup \partial M$$

is called the ***closure*** of M.

Remark. The closure of M is a closed set. We note that int $M \subseteq M$ then int $M \cup \partial M \subseteq M \cup \partial M$. We also have int $(V\backslash M) \cap M = \emptyset$ then $M \subseteq \text{int } M \cup \partial M$. Then we have $M \cup \partial M = \text{int } M \cup \partial M$. Then $V \backslash (M \cup \partial M) = \text{int } (V \backslash M)$ which is an open set. We then have $M \cup \partial M$ is a closed set.

The closure of a set may also be characterized in terms of sequences:

Lemma 0.6. Let $(V, \|\cdot\|)$ be a normed vector space and $M \subset V$. Then

$$\overline{M} = \left\{ x \in M : \underset{(x_n)_{n\in\mathbb{N}}}{\exists} x_n \in M \text{ and } x_n \to x \right\} \tag{0.7}$$

Proof.

(i) Suppose that $x \in V$ is such that there exists a sequence (x_n) with $x_n \in M$ and $x_n \to x$. Then for every $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that $B_\varepsilon(x)$ contains x_n for $n > N$. Hence, $B_\varepsilon \cap M \neq \emptyset$ and so x can not be an exterior point (Not an interior point of $V \backslash M$). This implies $x \in M \cup \partial M$.

(ii) Suppose $x \in M \partial M$. Then for every $\varepsilon > 0$, $B_\varepsilon(x) \cap M \neq \emptyset$ ($x \in B_\varepsilon(x)$ for $x \in M$ and from definition of $x \in \partial M$). Choose $\varepsilon = 1/n$ for $n \in \mathbb{N} \backslash \{0\}$ to find a sequence of points $x_n \in B_{1/n}(x) \cap M$. This sequence converges to x, so x is in the set on the right-hand side of (0.7).

□

Recall the following definition of continuity in normed vector spaces:

Definition 0.8. Let $(U, \|\cdot\|_1)$ and $(V, \|\cdot\|_2)$ be normed vector spaces and $f : U \to V$ a function. Then f is ***continuous at*** $a \in U$ if

$$\underset{\varepsilon>0}{\forall} \underset{\delta>0}{\exists} \underset{x\in U}{\forall} \|x - a\|_1 < \delta \Rightarrow \|f(x) - f(a)\|_2 < \varepsilon \tag{0.8}$$

Of course, we can prove as usual the following:

Theorem 0.7. Let $(U, \|\cdot\|_1)$ and $(V, \|\cdot\|_2)$ be normed vector spaces and $f : U \to V$ a function. Then f is ***continuous at*** $a \in U$ if and only if

$$\underset{\substack{(x_n)_{n\in\mathbb{N}} \\ x_n \in U}}{\forall} x_n \to a \Rightarrow f(x_n) \to f(a). \tag{0.9}$$

Proof. From definition of continuous function, we have for any $\varepsilon > 0$, there exists $\delta > 0$ such that

$$|x - a| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon.$$

Because $x_n \to a$, for $\delta > 0$, we can find $N > 0$ such that

$$\underset{n>N}{\forall} |x_n - a| < \delta,$$

which means

$$\underset{n>N}{\forall} |f(x_n) - f(a)| < \varepsilon.$$

Then $(f(x_n))$ converges to $f(a)$. □

Suppose that $f : M \to N$, where M, N are any sets. Let $A \subset M$. Then we define the ***image of A*** by

$$f(A) := \{y \in N : y = f(a) \text{ for some } x \in A\}.$$

In particular, we can write

$$\operatorname{ran} f = f(M).$$

Similarly, for $B \subset M$ we define the ***pre-image of B*** by

$$f^{(-1)}(B) := \{x \in M : f(x) = y \text{ for some } y \in B\}. \tag{0.10}$$

Example. (i) Let $f : \mathbb{R} \to \mathbb{R}$, $f(x) = \sin x$. Then $f([0, \pi]) = [0, 1]$.

(ii) Let $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^2 + y^2$. Then

$$f^{-1}(1) = \big\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\big\}$$

(This is a unit circle in $\mathbb{R}^2$).

It is often useful to characterize continuous maps by using open sets:

Theorem 0.8. Let $(U, \|\cdot\|_1)$ and $(V, \|\cdot\|_2)$ be normed vector spaces and $f : U \to V$ a function. Then f is continuous if and only if the pre-image $f^{-1}(\Omega)$ of every open set $\Omega \subset V$ is open.

Why we cannot have $f(\Omega)$ is open for $\Omega \subset U$ that is open?

Proof. ($\Rightarrow$) Let f be continuous and $\Omega \subset V$ open. We will show that $f^{-1}(\Omega)$ is open. Let $a \in f^{-1}(\Omega)$. Then $f(a) \in \Omega$, and since Ω is open we can find $\varepsilon > 0$ such that $B_\varepsilon(f(a)) \subset \Omega$.
Now let $\delta > 0$. By the continuity of f we can choose δ small enough to ensure that $f(B_\delta(a)) \subset B_\varepsilon(f(a))$. But then $B_\delta(a) \subset f^{-1}(\Omega)$. Since this is true for any $a \in f^{-1}(\Omega)$, it follows that $f^{-1}(\Omega)$ is open.
($\Leftarrow$) Let $f : U \to V$ be such that the pre-image $f^{-1}(\Omega)$ of every open set Ω is open. We will show that f is continuous. Let $a \in U$ be arbitrary and fix $\varepsilon > 0$. We want to show that there exists a $\delta > 0$ such that

$$x \in B_\delta(a) \Rightarrow f(x) \in B_\varepsilon(f(a)). \tag{0.11}$$

The set $B_\varepsilon(f(a))$ is open, and by assumption $f^{-1}(B_\varepsilon(f(a))) \ni a$ is also open. Thus, we can find $\delta > 0$ such that $B_\delta(a) \subset f^{-1}(B_\varepsilon(f(a)))$. But then we can find $\delta > 0$ for any $\varepsilon > 0$. This means f is a continuous function. $\square$

Example. We show that the function

$$\det : \operatorname{Mat}(n \times n; \mathbb{C}) \to \mathbb{C}, \quad \det A = \sum_{\pi \in S_n} \operatorname{sgn} \pi \; a_{\pi(1)1} \cdots a_{\pi(n)n}$$

is continuous.
In particular, we can choose to use the norm $\|A\| = \max_{i,j} |a_{ij}|$. Then fix $A = (a_{ij}) \in \operatorname{Mat}(n \times n; \mathbb{C})$ and suppose that (A_m) is a sequence converging to A. Our choice of norm implies that all coefficients converge, $a_{ij}^{(m)} \to a_{ij}$. Since $\det A$ is a polynomial in the coefficients a_{ij}, $\det A_m \to \det A$ and therefore det is continuous at $A \in \operatorname{Mat}(n \times n; \mathbb{C})$.
Note that the pre-image of the set of non-zero complex numbers is

$$\det{}^{-1}(\mathbb{C} \backslash \{0\}) = \operatorname{GL}(n; \mathbb{C}),$$

the ***general linear*** group of invertible matrices. Since $\mathbb{C}\setminus\{0\}$ is an open set, Theorem 0.8 implies that $\mathrm{GL}(n;\mathbb{C})$ is an open set in $\mathrm{Mat}(n \times n;\mathbb{C})$.

We are now interested in generalizing the results of Vv186 that apply to continuous functions on closed intervals to vector spaces. Note that a closed interval in $\mathbb{R}$ is always bounded in the following sense

Definition 0.9. Let $(V, \|\cdot\|)$ be a normed vector space and $M \subset V$. Then M is said to be ***bounded*** if there exists some $R > 0$ such that $M \subset B_R(0)$.

It turns out that the natural generalization of a closed interval is a little more complicated than just requiring a set to be closed and bounded.

Definition 0.10. Let $(V, \|\cdot\|)$ be a normed vector space and $K \subset V$. Then K is said to be ***compact*** if every sequence in K has a convergent subsequence with limit contained in K.

Theorem 0.9. Let $(V, \|\cdot\|)$ be a (possibly infinite-dimensional) normed vector space and $K \subset V$. Then K is closed and bounded.

Proof. We first show that K is closed by establishing $K = \overline{K}$. Let $x \in \overline{K}$. Then there exists a sequence (x_n) in K converging to x. Since K is compact, (x_n) has a subsequence (x_{n_k}) that converges to $x' \in K$. Since (x_n) converges to x, $x = x' \in K$, so $\overline{K} \subseteq K$. From definition, $K \subseteq \overline{K}$, so $K = \overline{K}$ and K is closed.
Now suppose that K is unbounded. Then for any $n \in \mathbb{N}$ there exists an $x_n \in K$ such that $\|x_n\| > n$. This gives rise to an unbounded sequence (x_n). Furthermore, any subsequence of (x_n) is unbounded, because the norm of a subsequence $\|x_{n_k}\|$ is greater than n_k and will always diverge. Since a convergent sequence is bounded, we conclude that (x_n) can not have a convergent subsequence. This implies that K is not compact. By contraposition, if K is compact, then K must be bounded. $\square$

Theorem 0.10. Let $(V, \|\cdot\|)$ be a ***finite-dimensional*** vector space and let $K \subset V$ be closed and bounded. Then K is compact.

Proof. Suppose that $(b_1, ..., b_n)$ be a basis of V and K closed and compact. Let (v_m) be a sequence in K. Then each sequence term has the representation

$$v_m = \lambda_1^{(m)} b_1 + \cdots + \lambda_n^{(m)} b_n, \qquad \lambda_1^{(m)}, ..., \lambda_n^{(m)} \in \mathbb{F}, \quad m \in \mathbb{N}.$$

By Lemma 0.4 and the boundedness of K, there exists constants $C_1, C_2 > 0$ such that

$$C_1 \geq \|v_m\| \geq C_2 \sum_{k=1}^{n} |\lambda_k^{(m)}|.$$

It follows that for each k, the sequence $\left(\lambda_k^{(m)}\right)$ is bounded. Write

$$\lambda^{(m)} = \left(\lambda_1^{(m)}, ..., \lambda_n^{(m)}\right).$$

By the Theorem of Bolzano-Weierstraß in $\mathbb{R}^n$, $\left(\lambda^{(m)}\right)$ has a convergent subsequence $\left(\lambda^{(m_j)}\right)$ so that (v_{m_j}) converges to some element $v \in \overline{K}$. Since K is closed, $v \in K$. This implies that K is compact. $\square$

We summarize this proof :

(i) since this set is bounded, any sequence has a convergent subsequence;

(ii) since this set is closed, any convergent sequence converges to a point in M.

Theorem 0.10 is in general false in infinite-dimensional spaces:

Example. Consider the vector space of summable complex sequences,

$$I^1 := \left\{ (a_n) : \mathbb{N} \to \mathbb{C} : \sum_{n=0}^{\infty} |a_n| < \infty \right\}.$$

The natural norm is given by

$$\|(a_n)\|_1 := \sum_{n=0}^{\infty} |a_n|.$$

Then the set

$$\overline{B_1(0)} = \left\{ (a_n) \in I^1 : \sum_{n=0}^{\infty} |a_n| \leq 1 \right\}$$

is closed and bounded, but not compact.

Why are we so interested in compact sets? Well, it turns out that compact sets are natural extensions of closed intervals in $\mathbb{R}$ for the purpose of generalizing some major theorems on continuous functions.

Theorem 0.11. Let $(U, \|\cdot\|_1)$, $(V, \|\cdot\|_2)$ be normed vector spaces and $K \subset U$ compact. Let $f : K \to V$ be continuous. Then $\operatorname{ran} f = f(K)$ is compact in V.

Proof. Let (y_n) be a sequence in $f(K)$. Then there exists a sequence (x_n) in K with $y_n = f(x_n)$. Since K is compact, a subsequence (x_{n_k}) of (x_n) converges to some $a \in K$. But because f is continuous the subsequence $(f(x_{n_k}))$ of (y_n) converges to some $f(a) \in f(K)$. Hence, (y_n) has a convergent subsequence and $f(K)$ is compact. $\square$

Remark. We note the key point is that a continuous function will map a convergent sequence to another convergent sequence.

Theorem 0.12. Let $(V, \|\cdot\|)$ be a normed vector space and $K \subset V$ compact. Let $f : K \to \mathbb{R}$ be continuous. Then f has a maximum in K, i.e., there exists an $x \in K$ such that $f(y) \leq f(a)$ for all $y \in K$.

Proof. The range $\operatorname{ran} f = f(K)$ is compact by Theorem 0.11, so it is closed and bounded by Theorem 0.9. The lease upper bound $b = \sup f(K)$ exists because $f(K)$ is bounded.
Since b is the ***least*** upper bound, b can not be en exterior point of $f(K)$, so $b \in \overline{f(K)}$.
Since $f(K)$ is closed, $\overline{f(K)} = f(K)$ and $b \in f(K)$.
Hence, there exists an $x \in K$ with $f(x) = b$ and $f(y) \leq b$ for all $y \in K$. $\square$

Remark. We recall that every bounded subset of real numbers has a least upper bound is $P13$ in axiom of real numbers.

Definition 0.11. Let $(U, \|\cdot\|_1)$ and $(V, \|\cdot\|_2)$ be normed vector spaces, $\Omega \subset U$ and $f : \Omega \to V$ a function. Then f is ***uniformly continuous in*** Ω if

$$\underset{\varepsilon>0}{\forall}\ \underset{\delta>0}{\exists}\ \underset{x,y\in\Omega}{\forall}\ \|x-y\|_1 < \delta \Rightarrow \|f(x) - f(y)\|_2 \varepsilon. \tag{0.12}$$

(Compare with Definition 0.8)

Theorem 0.13. Let $(U, \|\cdot\|_1)$ and $(V, \|\cdot\|_2)$ be normed vector spaces, $K \subset U$ a compact set and $f : K \to V$ continuous on K. Then f is uniformly continuous on K.

Proof. Suppose that f is continuous but not uniformly continuous on K. Then

$$\underset{\varepsilon>0}{\exists}\ \underset{\delta>0}{\forall}\ \underset{x,y\in K}{\exists} \|x-y\|_1 < \delta \wedge \|f(x)-f(y)\|_2 \geq \varepsilon.$$

Denote this ε by ε_0. Then for each $\delta = 1/n$ there exists vectors $x_n, y_n \in K$ such that

$$\|x_n - y_n\|_1 < \frac{1}{n} \wedge \|f(x) - f(y)\|_2 \geq \varepsilon_0.$$

Since K is compact, we can consider (x_n, y_n) as a sequence of U^2 vectors. From Theorem of Bolzano Weierstraß, we have there exists a subsequence (x_{n_k}, y_{n_k}) converges to (ξ, η). Since $\|x_{n_k} - y_{n_k}\| \leq \frac{1}{n_k}$, we see that $\xi = \eta$. However, then

$$x_{n_k} \to \xi \wedge y_{n_k} \to \xi \wedge \|f(x_{n_k}) - f(y_{n_k})\|_2 \geq \varepsilon_0 \nrightarrow 0.$$

Since K is compact, we have $\xi \in K$. Because f is continuous, because $x_{n_k} \to \xi$, $f(x_{n_k}) \to f(\xi)$, because $y_{n_k} \to \xi$, $f(y_{n_k}) \to f(\xi)$. This contradicts $\|f(x_{n_k}) - f(y_{n_k})\|_2 \geq \varepsilon_0$. $\square$

Remark. Note that K is compact guarantees that f is continuous at ξ.

Functions and Derivatives

PART II

To represent a function $f : \mathbb{R}^2 \to \mathbb{R}$, i.e., a real function of two variables. One method is using a three-dimensional graph showing and (x_1, x_2, z)-axes and plotting $z = f(x_1, x_2)$.

Another representation for functions $f : \mathbb{R}^2 \to \mathbb{R}$ is so-called ***contour plot***. In this two-dimensional graph we plot curves

$$C_\alpha = f^{-1}(\{\alpha\})$$

for several values of α. These are the pre-image sets of $\{\alpha\}$.

In the hamiltonian formulation of analytical mechanics, one defined a so-called ***Hamilton function*** H for a mechanical system. We have

$$H = T + V$$

where T is kinetic energy and V is potential energy. H remains constant if the system satisfies the law of energy conservation.

In the hamiltonian formulation of mechanics, the essential variables of a system are the position x and the momentum p. The variables are tracked in a so-called ***phase space*** $\mathbb{R}^n_x \times \mathbb{R}^n_p = \mathbb{R}^{2n}_{(x,p)}$, where typically, $n = 1, 2$ or 3. The time-evolution of the system is represented through ***phase curves*** in $\mathbb{R}^{2n}$, which are given by the contour lines of H, which is regarded as a function $\mathbb{R}^{2n} \to \mathbb{R}$. In other words, a phase curve is the set $H^{-1}(E)$, where E is the conserved energy of the system.

Example. For the simple harmonic oscillator, the kinetic energy is given by $T = \frac{1}{2}mv^2 = p^2/2m$ and the potential function is given by $V = \frac{k}{2}x^2$, so

$$H(x, p) = \frac{1}{2m}p^2 + \frac{k}{2}x^2.$$

In the rest of this term we will develop calculus for "functions of multiple variables". This generally means function defined on (a subset of) $\mathbb{R}^n$, but it is not any more difficult to treat functions defined on finite-dimensional vector spaces.

Throughout the following discussion, we assume that V and X denote finite-dimensional, normed vector spaces. The concrete norm will be irrelevant, as all norms are equivalent (see Theorem 0.1). We will consider first the derivative of a function

$$f : X \to V.$$

Definition 0.12. Let $f : X \to V_1$, $g : X \to V_2$ and $x_0 \in X$. We say that

$$f(x) = o(g(x)) \quad \text{as } x \to x_0 \quad \Leftrightarrow \quad \lim_{x \to x_0} \frac{\|f(x)\|_{V_1}}{\|g(x)\|_{V_2}} = 0.$$

Definition 0.13. Let X, V is finite-dimensional vector spaces and $\Omega \subset X$ an open set. Then a map $f : \Omega \to V$ is called ***differentiable at*** $x \in \Omega$ if there exists a liner map

$L_x \in \mathcal{L}(X, V)$ such that

$$f(x+h) = f(x) + L_x h + o(h) \qquad \text{as } h \to 0. \tag{0.13}$$

In this case we call L_x the ***derivative of*** f ***at*** x and write

$$L_x = Df|_x = df|_x.$$

We say that f is differentiable on Ω if it is differentiable for every $x \in \Omega$.

Remark.
- Just as in proof of 186 Lemma 3.1.2 we can show that the derivative is uniquely defined by ().
- We may also copy the proof of Lemma 186 3.1.8 to see that every differentiable function is continuous.

If f is differentiable on Ω, we may regard Df as a map

$$Df : \Omega \to \mathcal{L}(X, V), \qquad x \mapsto Df|_x.$$

Definition 0.14. We define

$$\begin{aligned} C(\Omega, V) &:= \{f : \Omega \to V : f \text{ is continuous}\}, \\ C^1(\Omega, V) &:= \{f : \Omega \to V : f \text{ is differentiable and } Df \text{ is continuous}\}. \end{aligned}$$

We may thus regard the ***derivative*** D as a (linear) map

$$D : C^1(\Omega, V) \to C(\Omega, \mathcal{L}(X, V)), \qquad f \mapsto Df.$$

Example. Let X, V be finite-dimensional vector spaces and $L \in \mathcal{L}(X, V)$ a linear map. Then

$$L(x+h) = Lx + Lh \overset{!}{=} Lx + DL|_x h + o(h) \qquad (h \to 0),$$

so the derivative of L at any $x \in X$ is $DL|_x = L$.

Example. Explicit instances of Example 1 are, e.g.,

- Let $X = V = \mathbb{C}$ be regarded as real vectors spaces and $f : z \to \overline{z}$ be the (then linear) complex conjugation. Then for $z, h \in \mathbb{C}$

  $$\overline{z+h} = \overline{z} + \overline{h},$$

  so $Df|_z(h) = \overline{h}$. Note Df is the complex conjugation.
- Regard $A \in \mathrm{Mat}(2 \times 2; \mathbb{R})$ as a linear map $\mathbb{R}^2 \to \mathbb{R}^2$. Then for $x, h \in \mathbb{R}^2$

  $$A(x+h) = Ax + Ah,$$

  so $DA|_x(h) = Ah$.
- Let $tr : \mathrm{Mat}(n \times n; \mathbb{C}) \to \mathbb{C}$ be the ***trace*** of a squared matrix, i.e.,

  $$\mathrm{tr}\, A = \mathrm{tr}(a_{ij})_{1 \le i,j \le n} = \sum_{i=1}^{n} a_{ii}.$$

Then the trace is linear and for $A, H \in \mathrm{Mat}(n \times n; \mathbb{C})$

$$D\ \mathrm{tr}|_A H = \mathrm{tr}\, H.$$

Example. Some examples of derivatives of non-linear maps are as follows:

- Let $X = V = \mathbb{C}$ be regarded as real vector spaces and $f : z \to z^2$. Then for $z, h \in \mathbb{C}$
$$(z+h)^2 = z^2 + 2zh + h^2,$$
so $Df|_z(h) = 2zh$.

- Let $f : \mathbb{R}^2 \to \mathbb{R}$ be given by
$$f(x) = f(x_1, x_2) = x_1 + 2x_2 x_2 + x_2^2.$$
Then, for $h = (h_1, h_2) \in \mathbb{R}^2$ and $x \in \mathbb{R}^2$,
$$\begin{aligned} f(x+h) &= f(x_1 + h_1, x_2 + h_2) \\ &= x_1 + h_1 + 2(x_2 + h_2)(x_1 + h_1) + (x_2 + h_2)^2 \\ &= f(x) + h_1 + 2(h_2 x_1 + x_2 h_1 + x_2 h_2) + 2h_1 h_2 + h_2^2 \end{aligned}$$
In
$$f(x+h) = f(x) + \underbrace{h_1 + 2(h_2 x_1 + h_1 x_2 + h_2 x_2)}_{=:L_{(x_1,x_2)}h} + 2h_1 h_2 + h_2^2$$
the term $L_{(x_1,x_2)}$ is clearly linear in h, while
$$\lim_{h \to 0} \frac{\|2h_1 h_2\|_{\mathbb{R}}}{\|h\|_{\mathbb{R}^2}} = \lim_{h_1, h_2 \to 0} \frac{|2h_1 h_2|}{\sqrt{h_1^2 + h_2^2}} = 2 \lim_{h_1, h_2 \to 0} \frac{|h_2|}{\sqrt{1 + (h_2/h_1)^2}}.$$
Since $|h_2| \to 0$ as $h_2 \to 0$ and $1/\sqrt{1 + (h_2/h_1)^2}$ i bounded, we see that
$$\lim_{h \to 0} = \frac{\|2h_1 h_2\|_{\mathbb{R}}}{\|h\|_{\mathbb{R}^2}} = 0,$$
and so $2h_1 h_2 = o(h)$ as $h \to 0$. Similarly, we show that $h^2 = o(h)$, so we conclude
$$Df|_x h = (1 + 2x_2)h_1 + 2(x_1 + x_2)h_2.$$

Remark. Notice that we may express the derivative as a 1×2 matrix,

$$Df|_x h = (1 + 2x_2,\ 2(x_1 + x_2)) \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}.$$

This is of course not surprising; if $X = \mathbb{R}^n$ and $V = \mathbb{R}^m$, i.e., we are considering a function

$$f : \mathbb{R}^n \supset \Omega \to \mathbb{R}^m, \qquad f(x_1, ..., x_n) = \begin{pmatrix} f_1(x_1, ..., x_n) \\ \vdots \\ f_m(x_1, ..., x_n) \end{pmatrix},$$

then its derivative at $x \in \Omega$ (if exists) is

$$Df|_x \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) \simeq \mathrm{Mat}(m \times n; \mathbb{R}).$$

How to obtain this matrix? Denote by e_j the jth standard basis vector in $\mathbb{R}^n$ or $\mathbb{R}^m$. We now consider the columns of $Df|_x$, which are given by $Df|_x e_j$, $j = 1, ..., n$. Assuming that f is differentiable, for any $h \in R$, $x \in \mathbb{R}^n$ and $j = 1, ..., n$ we have

$$f(x + he_j) = f(x) + Df|_x(he_j) + o(h),$$

which we may rewrite as

$$Df|_x e_j = \frac{1}{h}(f(x + he_j) - f(x)) + o(1) = \frac{1}{h}\sum_{k=1}^{m}(f_k(x + he_j) - f_k(x))e_k + o(1).$$

The $(i, j)$the element of $Df|_x$ is given by $\langle\, e_i\,,\, Df|_x e_j\,\rangle$,

$$(Df|_x)_{ij} = \langle\, e_i\,,\, Df|_x e_j\,\rangle = \frac{1}{h}(f_i(x + he_j) - f_{i(x)}) + o(1).$$

We now take the limit $h \to 0$ to obtain

$$(Df|_x)_{ij} = \langle\, e_i\,,\, Df|_x e_j\,\rangle = \lim_{h\to 0}\frac{f_i(x + he_j) - f_i(x)}{h}.$$

This gives us an intuition on Df.

Definition 0.15. Let $\Omega \subset \mathbb{R}^n$ and $f : \Omega \to \mathbb{R}$ be differentiable on Ω. We then define the ***partial derivative with respect to*** x_j ***at*** $x \in \Omega$ by

$$\begin{aligned}\left.\frac{\partial f}{\partial x_j}\right|_x &:= \lim_{h\to 0}\frac{f(x + he_j) - f(x)}{h}\\ &= \lim_{h\to 0}\frac{f(x_1, ..., x_{j-1}, x_j + h, x_{j+1}, ..., x_n) - f(x)}{h}\end{aligned}$$

In this notation,

$$(Df|_x)_{ij} = \frac{\partial f_i}{\partial x_j}$$

or rather

$$Df|_x = \left.\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}\right|_x$$

Remark. There are several notations for the partial derivatives of a function. If $f : \mathbb{R}^n \to \mathbb{R}$, we may use any of the following:

$$\frac{\partial f}{\partial x_j} = \partial_{x_j} f = \partial_j f = f_{x_j} = f_j$$

to denote differentiation w.r.t. the variable x_j.
In practice, we calculate the partial derivative w.r.t. x_j by ***holding all other variables constant*** and simply differentiating f as a function of x_j.

Example. Let $f(x_1, x_2, x_3) = x_1 \sin(x_1, x_2, x_3) + 3x_2^2 x_1$. Then

$$\begin{aligned} \frac{\partial f}{\partial x_1} &= \sin(x_1 x_2 x_3) + x_1 x_2 x_3 \cos(x_1 x_2 x_3) + 3x_2^2, \\ \frac{\partial f}{\partial x_2} &= x_1^2 x_3 \cos(x_1 x_2 x_3) + 6x_2 x_1, \\ \frac{\partial f}{\partial x_3} &= x_1^2 x_2 \cos(x_1 x_2 x_3). \end{aligned}$$

We note that if $Df|_x$ exists, we may write it as a matrix of partial derivatives. However, it is not clear whether the existence of all partial derivatives implies the existence of the derivative $Df|_x$. Thus it is useful to consider the matrix of partial derivatives on its own; in fact, it deserves a special designation.

Definition 0.16. Let $\Omega \subset \mathbb{R}^n$ and $f : \Omega \to \mathbb{R}^m$. Assume that all partial derivatives $\frac{\partial f_i}{\partial x_j}$ of f exists at $x \in \Omega$. The matrix

$$J_f(x) := \left.\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}\right|_x$$

called the ***Jacobian*** of f.

Remark. If the derivative $Df|_x \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ exists, $J_f(x) \in \mathrm{Mat}(m \times n; \mathbb{R})$ is the representing matrix of $Df|_x$ w.r.t. the standard bases in $\mathbb{R}^n$ and $\mathbb{R}^m$.

Example. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be given by $f(x_1, x_2) = (x_1^2 + x_2^2, x_2 - x_1)$. Then the partial derivatives are

$$\begin{aligned} \frac{\partial f_1}{\partial x_1} &= \frac{\partial}{\partial x_1}(x_1^2 + x_2^2) = 2x_1, & \frac{\partial f_1}{\partial x_2} &= \frac{\partial}{\partial x_2}(x_1^2 + x_2^2) = 2x_2, \\ \frac{\partial f_2}{\partial x_1} &= \frac{\partial}{\partial x_1}(x_2 - x_1) = -1, & \frac{\partial f_2}{\partial x_2} &= \frac{\partial}{\partial x_2}(x_2 - x_1) = 1. \end{aligned}$$

The Jacobian is given by

$$J_f(x_1, x_2) = \begin{pmatrix} 2x_1 & 2x_2 \\ -1 & 1 \end{pmatrix}$$

A natural question that arises is, "Does the existence of $J_f(x)$ imply the differentiability of f at x?"

However, the answer is negative. The existence of partial derivative is not a condition strong enough to guarantee the exists of derivative.

Theorem 0.14. Let $\Omega \subset \mathbb{R}^n$ be an open set and $f : \Omega \to \mathbb{R}^m$ such that all partial derivatives $\partial_{x_j} f$ exist on Ω.

(i) If all partial derivatives are bounded (there exists a constant $M > 0$ such that $|\partial_{x_j} f_i| \leq M$ on Ω), then f is continuous i.e., $f \in C(\Omega, \mathbb{R}^m)$.

(ii) If all partial derivatives are continuous on Ω, then f is continuously differentiable on Ω, i.e., $f \in C^1(\Omega, \mathbb{R}^m)$. In particular,

$$Df|_x = J_f(x) = \left.\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}\right|_x$$

for all $x \in \Omega$.

Proof. Let

$$f : \mathbb{R}^n \to \mathbb{R}^m, \qquad f(x) = \begin{pmatrix} f_1(x_1, ..., x_n) \\ \vdots \\ f_m(x_1, ..., x_n) \end{pmatrix}.$$

Here we need an important equation to use properties of partial derivatives to make arguments on $f(x+h) - f(x)$. We that $n = 2$ and consider $f_i(x)$

$$\begin{aligned} f_i(x+h) - f_i(x) &= f_i(x_1 + h_1, x_2 + h_2) - f_i(x_1, x_2) \\ &= [f_i(x_1 + h_1, x_2 + h_2) - f_i(x_1 + h_1, x_2)] + [f_i(x_1 + h_1, x_2) - f_i(x_1, x_2)] \end{aligned}$$

We want to apply Mean Value Theorem 3.2.7 of Vv186, define

We need Ω is an open set such that $(x_1 + h_1, y) \in \Omega$ is possible for small enough h.

$$g : \mathbb{R} \to \mathbb{R}, \qquad g(y) = f_i(x_1 + h_1, y).$$

Then there exists a $\theta_2 \in (x_2, x_2 + h_2)$ such that

$$\begin{aligned} f_i(x_1 + h_1, x_2 + h_2) - f_i(x_1 + h_1, x_2) &= g(x_2 + h_2) - g(x_2) \\ &= h_2 \cdot g'(\theta_2) \\ &= h_2 \partial_2 f_i(x_1 + h_1, x_2 + \tau_2 h_2) \end{aligned}$$

where we have chosen $\tau_2 \in (0, 1)$ such that $\theta_2 = x_2 + \tau_2 h_2$.
Similarly, we find that

$$f_i(x_1 + h_1, x_2) - f_i(x_1, x_2) = h_1 \frac{\partial f_i}{\partial x_1}(x_1 + \tau_1 h_1, x_2)$$

for some $\tau_1 \in (0, 1)$. We then have

$$f_i(x+h) - f_i(x) = h_1 \frac{\partial f_i}{\partial x_1}(x_1 + \tau_1 h_1, x_2) + h_2 \frac{\partial f_i}{\partial x_2}(x_1 + h_1, x_2 + \tau_2 h_2).$$

Generalizing to $n \geq 2$, we ave constants $\tau_1, ..., \tau_n \in (0, 1)$ such that

$$\begin{aligned} f_i(x+h) - f_i(x) =& f_i(x_1 + h_1, x_2 + h_2, ..., x_n + h_n) - f_i(x_1, x_2 + h_2, ..., x_n + h_n) \\ &+ f_i(x_1, x_2 + h_2, ..., x_n + h_n) - f_i(x_1, x_2, x_3 + h_3, ..., x_n + h_n) \\ &+ \cdots + f_i(x_1, x_2, ..., x_{n-1}, x_n + h_n) - f_i(x_1, x_2, ..., x_n) \\ =& h_1 \partial_1 f_i(x_1 + \tau_1 h_1, x_2 + h_2, ..., x_n + h_n) \\ &+ h_2 \partial_2 f_i(x_1, x_2 + \tau_2 h_2, x_3 + h_3, ..., x_n + h_n) \\ &+ \cdots + h_n \partial_n f_i(x_1, x_2, ..., x_n + \tau_n h_n). \end{aligned}$$

We introduce notation

$$u^j = (x_1, ..., x_{j-1}, x_j + \tau_j h_j, x_{j+1} + h_{j+1}, ..., x_n + h_n)$$

for $j = 1, ..., n$.
We then can write the previous equation as

$$f_i(x+h) - f_i(x) = \sum_{j=1}^{n} h_j \partial_j f_i(u^j).$$

We proceed with the proof of the theorem.

(i) Suppose that the partial derivatives are bounded. We want ot prove that f is continuous at $x \in \Omega$, i.e.,

$$\lim_{h \to 0} f(x+h) = f(x)$$

where are are free to choose arbitrary norms in $\mathbb{R}^n$ and $\mathbb{R}^m$ for the convergence. In both spaces we choose the maximum norm $\|\cdot\|_\infty$:

$$\begin{aligned}
\|f(x+h) - f(x)\|_\infty &= \max_{i=1,\dots,m} |f_i(x+h) - f_i(x)| \\
&\leq m \cdot \max_{i=1,\dots,m} \Big| \sum_{j=1}^{n} h_j \partial_j f_i(u^j) \Big| \\
&\leq m \cdot \max_{i=1,\dots,m} n \cdot \max_{j=1,\dots,n} h_j \cdot \max_{\substack{i=1,\dots,m \\ j=1,\dots,n}} \sup_{x \in \Omega} \left| \frac{\partial f_i}{\partial x_j}(x) \right| \\
&\leq m \cdot n \cdot \|h\|_\infty \cdot M \xrightarrow{h \to 0} 0.
\end{aligned}$$

(ii) Write

$$L = \left(\frac{\partial f_i}{\partial x_j} \right)_{\substack{i=1,\dots,m \\ j=1,\dots,n}} = (L_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$$

for the Jacobian. We want to show that

$$f(x+h) - f(x) - Lh = o(h) \qquad \text{as } h \to 0.$$

We again choose the maximum norm $\|\cdot\|_\infty$ to establish the convergence. We have the following estimate:

$$\begin{aligned}
\|f(x+h) - f(x) - Lh\|_\infty &= \max_{i=1,\dots,m} \left| f_i(x+h) - f_i(x) - \sum_{j=1}^{n} L_{ij} h_j \right| \\
&= \max_{i=1,\dots,m} \left| \sum_{j=1}^{n} h_j (\partial_j f_i(u^j) - \partial_j f_i(x)) \right| \\
&\leq \|h\|_\infty \sum_{j=1}^{n} \underbrace{\max_{i=1,\dots,m} |\partial_j f_i(u^j) - \partial_j f_i(x)|}_{\to 0 \text{ as } h \to 0} \\
&= o(h) \qquad \text{as } h \to 0.
\end{aligned}$$

Observe that we use the assumption that $\partial_j f_i(x)$ is continuous as x. This proves that f is differentiable, $L = Df|_x$ and $Df|_x$ depends continuously on x.

$\square$

Remark. In this proof, we establish the connection between values of functions and its derivatives using Mean Value Theorem. Note we can only apply Mean Value Theorem now on one dimension, so it is necessary to decompose $f(x+h) - f(x)$ in previous form.
Let $\Omega \subset \mathbb{R}^n$ be an open set. Then

$$C^1(\Omega, \mathbb{R}^m) = \{f : \Omega \to \mathbb{R}^m : \partial_j f_i \text{ is continuous for } j = 1, ..., n \\ \text{and } i = 1, ..., m\}.$$

If $m = 1$, we write $C^1(\Omega) := C^1(\Omega, \mathbb{R})$ for short.

We now establish the product and chain rules for differentiation.

To avoid having to re-prove the product rule for various types of products that we will encounter, we first define a generalized product through those properties that we shall need.

Definition 0.17. Let X_1, X_2, V be normed vector spaces. A map $\odot : X_1 \times X_2 \to V$ is called a ***(generalized product)*** if

1. $\odot$ is bilinear, i.e., linear in each entry and
2. $\|u \odot v\|_V \leq \|u\|_{X_1} \|v\|_{X_2}$ for all $u \in X_1$, $v \in X_2$.

Example.
1. The scalar product in $\mathbb{R}^n$;
2. The cross product $\times : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$;
3. For a compact non-empty set $K \subset \mathbb{R}^n$ and $f, g \in C(K, \mathbb{R})$ the pointwise product $f \cdot g \in C(K, \mathbb{R})$, defined by

$$(f \cdot g)(x) = f(x)g(x)$$

Theorem 0.15. *(Product Rule)* Let U, X_1, X_2, V be finite-dimensional vector spaces and $\Omega \subset U$ and open set. Let $f : \Omega \to X_1$ ad $g : \Omega \to X_2$ be differentiable maps and $\odot : X_1 \times X_2 \to V$ a generalized product. Then $f \odot g : \Omega \to V$ is also differentiable and

$$D(f \odot g) = (Df) \odot g + f \odot (Dg). \tag{0.14}$$

At $x \in \Omega$ the right-hand side is interpreted as a linear map $U \to V$

$$u \mapsto D(f \odot g)|_x u = (Df|_x u) \odot g(x) + f(x) \odot (Dg|_x u). \tag{0.15}$$

Proof. The proof is similar ot that for the product rule for functions of one variable. We telescope the difference,

$$\begin{aligned} & f(x+h) \odot g(x+h) - f(x) \odot g(x) \\ =& f(x+h) \odot (g(x+h) - g(x)) + (f(x+h) - f(x)) \odot g(x) \\ =& (f(x) + O(h)) \odot (Dg|_x h + o(h)) + (Df|_x h + o(h)) \odot g(x) \end{aligned}$$

as $h \to 0$. Here we use $Df|_x + o(h) = O(h)$. Extending the relevant limit theorems form the pointwise product to the generalized product, we have

$$\begin{aligned}&f(x+h)\odot g(x+h)-f(x)\odot g(x)\\=&f(x)\odot(Dg|_x h)+O(\|h\|^2)+o(h)+(Df|_x h)\odot g(x)+o(h)\\=&f(x)\odot(Dg|_x h)+(Df|_x h)\odot g(x)+o(h)\end{aligned}$$

$\square$

Remark. To have $O(h)\odot(Dg|_x h+o(h))=O(\|h\|^2)$, we need the property of generalized product.

Theorem 0.16. *(Chain Rule)* Let U, X, V be finite-dimensional vector spaces and $\Omega\subset U$, $\Sigma\subset X$ open sets. Let $g:\Omega\to\Sigma$ and $f:\Sigma\to V$ ve differentiable maps. Then the composition $f\circ g:\Omega\to V$ is also differentiable and for all $x\in\Omega$

$$D(f\circ g)|_x = Df|_{g(x)}\circ Dg|_x, \tag{0.16}$$

where the right-hand side is a composition of linear maps.

Proof. We assume that g is differentiable at x we have

$$(f\circ g)(x+h)=f(g(x+h)+Dg|_x h+o(h))$$

as $h\to 0$. Writing $H=Dg|_x h+o(h)$ and using the differentiability of f at $g(x)$, we have

$$(f\circ g)(x+h)=f(g(x)+H)=f(g(x))+Df_{g(x)}H+o(H).$$

Now

$$H=Dg|_x h+o(h)=O(h)+O(h)=O(h)$$

as $h\to 0$, so $o(H)=o(O(h))=o(h)$ as $h\to 0$.
We hence have

$$\begin{aligned}(f\circ g)(x+h)&=f(g(x))+Df_{g(x)}H+o(h)\\&=f(g(x))+Df|_{g(x)}\circ(Dg|_x h+o(h))+o(h)\\&=f(g(x))+Df|_{g(x)}\circ Dg|_x h+o(h)\end{aligned}$$

as $h\to 0$, so $D(f\circ g)(x)=Df_{g(x)}\circ Df_x$. $\square$

Example. Consider the polar coordinates $(r,\phi)\in(0,\infty)\times[0,2\pi)$, defined through the map

$$\Phi(r,\phi)=\begin{pmatrix}r\cos\phi\\ r\sin\phi\end{pmatrix}.$$

Then

$$D\Phi|_{(r,\phi)}=\begin{pmatrix}\frac{\partial\Phi_1}{\partial r}&\frac{\partial\Phi_1}{\partial\phi}\\ \frac{\partial\Phi_2}{\partial r}&\frac{\partial\Phi_2}{\partial\phi}\end{pmatrix}=\begin{pmatrix}\cos\phi&-r\sin\phi\\ \sin\phi&r\cos\phi\end{pmatrix}.$$

Next, consider the map $U:\mathbb{R}^2\to\mathbb{R}$, $(x_1,x_2)\mapsto x_1^2+x_2^2$. The derivative is

$$DU|_x=\left(\frac{\partial U}{\partial x_1},\frac{\partial U}{\partial x_2}\right)=(2x_1,2x_2)$$

Now $U\circ\Phi=(r\cos\phi)^2+(r\sin\phi)^2=r^2$. Clearly, $D(U\circ\Phi)|_{(r,\phi)}=(2r,0)$. We can also apply the chain rule:

$$\begin{aligned}
D(U\circ\Phi)|_{(r,\phi)} &= DU|_{(r\cos\phi, r\sin\phi)} D\Phi|_{r,\phi} \\
&= (2r\cos\phi, 2r\sin\phi)\begin{pmatrix}\cos\phi & -r\sin\phi \\ \sin\phi & r\cos\phi\end{pmatrix} \\
&= (2r\cos^2\phi + 2r\sin^2\phi, -2r^2\cos\phi\sin\phi + 2r^2\sin\phi\cos\phi) \\
&= (2r, 0)
\end{aligned}$$

In this part, we first consider the integration of a vector-space-valued but defined on $I = [a, b] \subset \mathbb{R}$.

$$\int_a^b f(x)dx, \qquad \text{where } f : [a, b] \to V.$$

Fortunately, the procedure is completely analogous to that of functions $f : \mathbb{R} \to \mathbb{R}$, at least for the regulated integral: we define step functions on $[a, b]$ with respect to a partition P by setting them constant on sub-intervals of the partition:

Definition 0.18. A V be a real or complex vector space. A function $f : [a, b] \to V$ is called a ***step function with respect to a partition*** $P = (a_0, ..., a_n)$ if there exists elements $y_i \in V$ such that $f(t) = y_i$ whenever $a_{i-1} < t < a_i$, $i = 1, ..., n$. We denote the set of all step functions by $\mathrm{Step}([a, b], V)$.

Remark. Note values on the partition boundaries can be arbitrary.

Example. The map

$$f : [0, 1] \to \mathbb{R}^2, \qquad f(x) = \begin{cases} \begin{pmatrix}0 \\ 1/2\end{pmatrix} & 0 \le x < 1/2 \\ \begin{pmatrix}1 \\ 1\end{pmatrix} & x = 1/2 \\ \begin{pmatrix}2 \\ 0\end{pmatrix} & 1/2 < x \le 1 \end{cases}$$

is a step function.

Definition 0.19. Let $I \subset \mathbb{R}$ be an interval and $(V, \|\cdot\|_V)$ a normed vector space. We say that a function $f : I \to V$ is ***bounded*** if

$$\|f\|_\infty := \sup_{x\in I} \|f(x)\|_V < \infty. \tag{0.17}$$

The set of all bounded functions $f : I \to V$ is denoted $L^\infty(I, V)$.

Example. The map

$$f : \mathbb{R} \to \mathbb{R}^2, \qquad f(t) = \begin{pmatrix}\sin t \\ e^{-t^2}\end{pmatrix}$$

is bounded map. To see this, we endow $\mathbb{R}^2$ with the norm $\|x\|_1 := |x_1| + |x_2|$. (Since all norms in $\mathbb{R}^n$ are equivalent, we can tak a convenient norm). Then

$$\|f\|_\infty := \sup_{t\in\mathbb{R}} \|f(t)\|_1 = \sup_{t\in\mathbb{R}}(|\sin t| + |e^{-t^2}|)$$
$$\sup_{t\in\mathbb{R}} |\sin t| + \sup_{t\in\mathbb{R}} |e^{-t^2}| = 2 < \infty.$$

We then define the integral of a step function as before:

Theorem 0.17. Let $f : [a, b] \to V$ be a step function with respect ot some partition P. Then

$$I_p(f) := (a_1 - a_0)y_1 + \cdots + (a_n - a_{n-1})y_n \in V$$

is independent of the choice of the partition P and is called the ***integral*** of f.

Partition P requires f is a step function with respect to P.

(This makes it impossible to define the Riemann integral for functions $f : I \to V$, because it relies on comparing the size of upper and lower step functions)

The main ingredient is again uniform, where we now say the a sequence of functions $(f_n), f_n : I \to V$, $I \subset \mathbb{R}$, converges uniformly to $f : I \to V$ in a normed vector space $(V, \|\cdot\|_V)$ if

$$\|f_n - f\|_\infty := \sup_{x\in I} \|f_n(x) - f(x)\|_V \xrightarrow{n\to\infty} 0.$$

A function f is them said to be ***regulated*** if it is the uniform limit of a sequence of step functions. We can then define the integral of f as the limit of the integrals of these step functions.

The upshot is the following: if $f : [a, b] \to \mathbb{R}^n$ is piecewise continuous, then f is regulated and

$$\int_a^b f(x)dx = \int_a^b \begin{pmatrix} f_1(x) \\ \vdots \\ f_n(x) \end{pmatrix} dx = \begin{pmatrix} \int_a^b f_1(x)dx \\ \vdots \\ \int_a^b f_n(x)dx \end{pmatrix}$$

(This follows because a sequence of step functions converging uniformly to f will converge uniformly in each component; the individual components are then equal to the "usual" regulated integrals of real-valued functions.)

Furthermore, we have the standard estimate

$$\left\|\int_a^b f(x)dx\right\|_V \le \int_a^b \|f(x)\|_V dx \le |b - a| \cdot \sup_{x\in[a,b]} \|f(x)\|_V.$$

Theorem 0.18. *Mean Value Theorem* Let (X, V) be finite-dimensional vector spaces, $\Omega \subset X$ open and $f \in C^1(\Omega, V)$. Let $x, y \in \Omega$ and assume that the line segment $x+ty$, $0 \le t \le 1$, is wholly contained in Ω. Then

$$f(x + y) - f(x) = \int_0^1 Df|_{x+ty}\, y\, dt = \left(\int_0^1 Df|_{x+ty}dt\right) y. \tag{0.18}$$

Remark. Note the second term is the integral of elements in V and the third term is the integral of elements in $\mathcal{L}(X, V)$.

Proof. Define the auxiliary function $g \in C^1([0, 1], V)$ by $g(t) := f(x + ty)$. Thus (by 186 Lemma 4.2.3) we have

$$f(x+y)-f(x)=g(1)-g(0)=\int_0^1 g'(t)dt.$$

For $\gamma(t)=x+ty$ we have $\gamma'(t)=y$. Applying the chain rule,

$$g'(t)=D(f\circ\gamma)|_t=Df|_{\gamma(t)}D\gamma|_t=Df_{x+ty}y.$$

Thus we obtain

$$f(x+y)-f(x)=\int_0^1 Df|_{x+ty}ydt,$$

proving the first inequality.
We now prove that y may be "taken out" of the integral. We want to use the linearity of integral. If we take the upper limit of integral as a parameter $z\in(0,1)$, we note that the derivatives of two sides with respect to this parameter are the same

$$\frac{d}{dx}\int_0^z Df|_{x+yt}\,y\,dt=\frac{d}{dz}\left\{\left(\int_0^z Df|_{x+yt}\right)y\right\}.$$

Furthermore, we have when $z=0$, both sides are equal to 0,

$$\int_0^0 Df|_{x+yt}\,y\,dt=0=\left(\int_0^0 Df|_{x+yt}dt\right)y.$$

Therefore,

$$\int_0^z Df|_{x+yt}\,y\,dt=\left(\int_0^z Df|_{x+yt}dt\right)y$$

for all $z\in[0,1]$, in particular for $z=1$. □

Remark. The Mean Value Theorem yields

$$\|f(x+y)-f(x)\|_V\le\|y\|_X\cdot\sup_{0\le t\le 1}\|Df|_{x+ty}\|,$$

where $Df|_{x+ty}$ denotes the operator norm of $Df|_{x+ty}\in\mathcal{L}(X,V)$.

Theorem 0.19. Let X,V be finite dimensional vector spaces, $I=[a,b]\subset\mathbb{R}$ an interval and $\Omega\subset X$ an open set. Let $f: I\times\Omega\to V$ be a continuous function such that $Df(t,\cdot)|_x$ exists and is continuous at every $(t,x)\in I\times\Omega$. Then

$$g(x)=\int_a^b f(t,x)dt$$

is differentiable in Ω and

$$Dg(x)=\int_a^b Df(t,\cdot)|_x dt.$$

Proof. Fix $x\in\Omega$ and choose h small enough such that $x+h\in\Omega$. In any case, we assume $\|h\|<1$. We want to find

$$g(x+h)=g(x)+Lh+o(h). \tag{0.19}$$

We rewrite $g(x+h)-g(x)$ and have

$$g(x+h)-g(x) = \int_a^b f(t,x+h)dt - \int_a^b f(t,x)dt$$
$$\int_a^b \left(f(t,x+h)-f(t,x)\right)dt$$

By the Mean Value Theorem, we have

$$\begin{aligned} g(x+h)-g(x) &= \int_a^b \left(\int_0^1 Df(t,\cdot)|_{x+sh} ds\right) dt\ h \\ &= \int_a^b \left(\int_0^1 Df(t,\cdot)|_x ds\right) dt\ h + \int_a^b \left(\int_0^1 Df(t,\cdot)|_{x+sh} - Df(t,\cdot)|_x ds\right) dt\ h. \end{aligned}$$

Here we add and subtract one $\int_a^b \left(\int_0^1 Df(t,\cdot)|_x ds\right) dt\ h = \int_a^b Df(t,\cdot)|_x dt\ h$. We want to show the last part is $o(h)$. We use the standard estimate on integral and have

$$\begin{aligned} \int_a^b \left(\int_0^1 Df(t,\cdot)|_{x+sh} - Df(t,\cdot)|_x ds\right) dt\ h &\leq (b-a) \sup_{t\in[a,b]} \left\| \int_0^1 (Df(t,\cdot)|_{x+sh} - Df(t,\cdot)|_x) ds \right\| \cdot \|h\|_X \\ &\leq (b-a) \sup_{t\in[a,b]} \sup_{s\in[0,1]} \|Df(t,\cdot)|_{x+sh} - Df(t,\cdot)|_x\| \cdot \|h\|_X \end{aligned}$$

Here $\|\cdot\|$ on $Df(t,\cdot)|_{x+sh} - Df(t,\cdot)|x$ denotes the operator norm. We now want to show that

$$\sup_{t\in[a,b]} \sup_{s\in[0,1]} \|Df(t,\cdot)|_{x+sh} - Df(t,\cdot)|_x\|$$

vanishes when $h\to 0$. To prove this, we need the continuity of $Df(t,\cdot)$.
Consider the function $Df(t,\cdot)|_y$ where y varies in the closed and bounded set $\overline{B_1(x)}$. Here we use $\|h\| < 1$. Since V is finite-dimensional, thi set is compact and so is $[a,b]\times\overline{B_1(x)}$.
Since $Df(t,\cdot)|_y$ is continuous in the compact set $[a,b]\times\overline{B_1(x)}$, it is also uniformly continuous by Theorem 0.13. Hence, we have when $\|h\|\to 0$, $\|Df(t,\cdot)|_{x+sh} - Df(t,\cdot)|_x\| \to 0$, for $(t, x+sh)\in[a,b]\times\overline{B_1(x)}$.
Then we have

$$g(x+h) = g(x) + \left(\int_a^b Df(t,\cdot)|_x dt\right) h + o(h).$$

□

For completeness, we attach the Inverse Function Theorem as follows:

Theorem 0.20. Suppose that $f:\mathbb{R}^n\to\mathbb{R}^n$ is continuously differentiable in an open set containing a, and $\det Df_{x=a}\neq 0$. Then there is an open set V containing a and an open set W containing $f(a)$ such that $f: V\to W$ has a continuous inverse $f^{-1}: W\to V$ which is differentiable and for any $y\in W$ satisfies

$$Df^{-1}|_y = \left[Df|_{f^{-1}(y)}\right]^{-1}.$$

Proof. The major difficulty in the proof is to show f^{-1} exists, continuous and differentiable. To have an intuition, we assume these conditions can be satisfied. We need to find the derivative :

$$f^{-1}(y+h) = f^{-1}(y) + Df^{-1}|_y h + o(h).$$

We apply f on both sides and

$$\begin{aligned} y + h &= f(f^{-1}(y) + Df^{-1}|_y h + o(h)) \\ &= y + Df|_{f^{-1}(y)}(Df^{-1}|_y h + o(h)) + o(h) \\ &= y + Df|_{f^{-1}(y)} \circ Df^{-1}|_y h + o(h) \end{aligned}$$

To have the equation holds when $h \to 0$, we need

$$Df|_{f^{-1}(y)} \circ Df^{-1}|_y = I,$$

and

$$Df^{-1}|_y = \left[Df|_{f^{-1}(y)}\right]^{-1}.$$

$\square$

Curves in Vector Spaces

PART

III

An important case of a map between vector space is a map $\mathbb{R} \to V$, where $(V, \|\cdot\|)$ is a normed vector space.

Definition 0.20. Let V be a finite-dimensional vector space and $I \subset \mathbb{R}$ and interval.

- A set $\mathcal{C} \subset V$ for which there exists a continuous, surjective and locally injective map $\gamma : I \to \mathcal{C}$ is called a ***curve***.
- The map γ is called a ***parametrization*** of $\mathcal{C}$.
- A curve $\mathcal{C}$ together with a parametrization is called a ***parametrized curve***.

Remark. Here ***locally injective*** means that in the neighborhood $B_\varepsilon(x) \cap I$ of any point $x \in I$ the parametrization is injective.
More generally, let f be a function of a single real variable. We say that a property holds ***locally at point*** $p \in \mathbb{R}$ if this property holds in some ε-neighborhood $B_\varepsilon(p)$.

Example. The set

$$S^1 := \left\{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 = 1\right\}$$

is a curve in $\mathbb{R}^2$ because we can find a parametrization, e.g.,

$$\gamma : [0, 2\pi] \to S^1, \qquad \gamma(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}.$$

It is clear that γ is continuous. The map γ is not injective, since $\gamma(0) = \gamma(2\pi) = (1, 0)$. But it is injective on $(0, 2\pi)$, we can find $\varepsilon = \pi$ such that γ is injective on $B_\varepsilon(x) \cap [0, 2\pi]$.

Definition 0.21. Let $\mathcal{C} \subset V$ be a curve possessing a parametrization $\gamma : I \to \mathcal{C}$ with int $I = (a, b)$ for $-\infty \leq a < b \leq \infty$.

(i) If γ is (globally) injective parametrization we say that $\mathcal{C}$ is a ***simple curve***.

(ii) If

$$\lim_{t \to a} \gamma(t) = \lim_{t \to b} \gamma(t),$$

the curve $\mathcal{C}$ is said to be ***closed***.

(iii) If a curve is not closed, it is said to be ***open***. Th point

$$x := \lim_{t \to a} \gamma(t) \qquad \text{and} \qquad y := \lim_{t \to b} \gamma(t)$$

are called the ***initial point*** and the ***final point*** of the parametrized curve $(\mathcal{C}, \gamma)$. The open curve is said to join x and y.

Remark. Whether a point is an initial point or final point of an open curve depends on the parametrization.

Example. The simple open curve

$$\mathcal{C} = \left\{(x_1, x_2) \in \mathbb{R}^2 : 0 \leq x_1 \leq 1, x_2 = x_1^2\right\}$$

joins the points $x = (0, 0)$ and $y = (1, 1)$. Either may be considered the initial point or the final point. Possible parametrizations are

$$\gamma(t) = \begin{pmatrix} t \\ t^2 \end{pmatrix}, \qquad \tilde{\gamma}(t) = \begin{pmatrix} 1 - t \\ (1 - t)^2 \end{pmatrix}$$

where both $\gamma, \tilde{\gamma} : [0, 1] \to \mathcal{C}$.

Potential Functions

PART

IV

The Second Derivative

PART

V

Extrema of Potential Functions

PART

VI

Constrained Extrema

PART

VII