Fundamentals of Linear Algebra and Optimization

CIS515
Part II: Optimization and Machine Learning. Some Slides

Jean Gallier
Department of Computer and Information Science
University of Pennsylvania
Philadelphia, PA 19104, USA
e-mail: jean@cis.upenn.edu

April 6, 2022

Chapter 1

Topology

1.1 Metric Spaces and Normed Vector Spaces

Most spaces that we consider have a topological structure given by a metric or a norm, and we first review these notions.

We begin with metric spaces.

Recall that $\mathbb{R}_+ = \{x \in \mathbb{R} \mid x \geq 0\}$.

Definition 1.1. *A metric space is a set E together with a function $d\colon E \times E \to \mathbb{R}_+$, called a metric, or distance, assigning a nonnegative real number $d(x, y)$ to any two points $x, y \in E$, and satisfying the following conditions for all $x, y, z \in E$:*

(D1) $d(x, y) = d(y, x)$. *(symmetry)*

(D2) $d(x, y) \geq 0$, *and* $d(x, y) = 0$ *iff* $x = y$. *(positivity)*

(D3) $d(x, z) \leq d(x, y) + d(y, z)$. *(triangle inequality)*

Geometrically, condition (D3) expresses the fact that in a triangle with vertices x, y, z, the length of any side is bounded by the sum of the lengths of the other two sides.

From (D3), we immediately get

$$|d(x, y) - d(y, z)| \leq d(x, z).$$

Let us give some examples of metric spaces. Recall that the *absolute value* $|x|$ of a real number $x \in \mathbb{R}$ is defined such that $|x| = x$ if $x \geq 0$, $|x| = -x$ if $x < 0$, and for a complex number $x = a + ib$, by $|x| = \sqrt{a^2 + b^2}$.

Example 1.1.

1. *Let $E = \mathbb{R}$, and $d(x, y) = |x - y|$, the absolute value of $x - y$. This is the so-called natural metric on $\mathbb{R}$.*

2. *Let $E = \mathbb{R}^n$ (or $E = \mathbb{C}^n$). We have the Euclidean metric*

$$d_2(x, y) = \left(|x_1 - y_1|^2 + \cdots + |x_n - y_n|^2\right)^{\frac{1}{2}},$$

the distance between the points $(x_1, \ldots, x_n)$ and $(y_1, \ldots, y_n)$.

3. *For every set E, we can define the discrete metric, defined such that $d(x, y) = 1$ iff $x \neq y$, and $d(x, x) = 0$.*

4. *For any* $a, b \in \mathbb{R}$ *such that* $a < b$*, we define the following sets:*

$[a, b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}$, *(closed interval)*

$(a, b) = \{x \in \mathbb{R} \mid a < x < b\}$, *(open interval)*

$[a, b) = \{x \in \mathbb{R} \mid a \leq x < b\}$, *(interval closed on the left, open on the right)*

$(a, b] = \{x \in \mathbb{R} \mid a < x \leq b\}$, *(interval open on the left, closed on the right)*

Let $E = [a, b]$*, and* $d(x, y) = |x - y|$*. Then,* $([a, b], d)$ *is a metric space.*

We will need to define the notion of proximity in order to define convergence of limits and continuity of functions. For this, we introduce some standard "small neighborhoods."

Definition 1.2. *Given a metric space E with metric d, for every $a \in E$, for every $\rho \in \mathbb{R}$, with $\rho > 0$, the set*

$$B(a, \rho) = \{x \in E \mid d(a, x) \leq \rho\}$$

is called the closed ball of center a and radius ρ, the set

$$B_0(a, \rho) = \{x \in E \mid d(a, x) < \rho\}$$

is called the open ball of center a and radius ρ, and the set

$$S(a, \rho) = \{x \in E \mid d(a, x) = \rho\}$$

is called the sphere of center a and radius ρ. It should be noted that ρ is finite (i.e., not $+\infty$). A subset X of a metric space E is bounded if there is a closed ball $B(a, \rho)$ such that $X \subseteq B(a, \rho)$.

Clearly, $B(a, \rho) = B_0(a, \rho) \cup S(a, \rho)$.

Example 1.2.

1. *In* $E = \mathbb{R}$ *with the distance* $|x - y|$*, an open ball of center* a *and radius* ρ *is the open interval* $(a - \rho, a + \rho)$*.*
2. *In* $E = \mathbb{R}^2$ *with the Euclidean metric, an open ball of center* a *and radius* ρ *is the set of points inside the disk of center* a *and radius* ρ*, excluding the boundary points on the circle.*
3. *In* $E = \mathbb{R}^3$ *with the Euclidean metric, an open ball of center* a *and radius* ρ *is the set of points inside the sphere of center* a *and radius* ρ*, excluding the boundary points on the sphere.*

One should be aware that intuition can be misleading in forming a geometric image of a closed (or open) ball.

For example, if d is the discrete metric, a closed ball of center a and radius $\rho < 1$ consists only of its center a, and a closed ball of center a and radius $\rho \geq 1$ consists of the entire space!

If $E = [a, b]$, and $d(x, y) = |x - y|$, as in Example 1.1, an open ball $B_0(a, \rho)$, with $\rho < b - a$, is in fact the interval $[a, a + \rho)$, which is closed on the left.

We now consider a very important special case of metric spaces, normed vector spaces.

Definition 1.3. *Let E be a vector space over a field K, where K is either the field $\mathbb{R}$ of reals, or the field $\mathbb{C}$ of complex numbers. A norm on E is a function $\| \; \|: E \to \mathbb{R}_+$, assigning a nonnegative real number $\|u\|$ to any vector $u \in E$, and satisfying the following conditions for all $x, y \in E$:*

(N1) $\|x\| \geq 0$, *and* $\|x\| = 0$ *iff* $x = 0$. *(positivity)*

(N2) $\|\lambda x\| = |\lambda| \, \|x\|$. *(scaling)*

(N3) $\|x + y\| \leq \|x\| + \|y\|$. *(convexity inequality)*

A vector space E together with a norm $\| \; \|$ is called a normed vector space.

From (N3), we easily get

$$|\|x\| - \|y\|| \leq \|x - y\|.$$

Given a normed vector space E, if we define d such that

$$d(x, y) = \|x - y\|,$$

it is easily seen that d is a metric. Thus, every normed vector space is immediately a metric space.

Note that the metric associated with a norm is invariant under translation, that is,

$$d(x + u, y + u) = d(x, y).$$

For this reason, we can restrict ourselves to open or closed balls of center 0.

Let us give some examples of normed vector spaces.

Example 1.3.

1. *Let* $E = \mathbb{R}$, *and* $\|x\| = |x|$, *the absolute value of* x. *The associated metric is* $|x - y|$, *as in Example 1.1.*

2. *Let* $E = \mathbb{R}^n$ *(or* $E = \mathbb{C}^n$*). There are three standard norms. For every* $(x_1, \ldots, x_n) \in E$, *we have the norm* $\|x\|_1$, *defined such that,*

$$\|x\|_1 = |x_1| + \cdots + |x_n|,$$

we have the Euclidean norm $\|x\|_2$, *defined such that,*

$$\|x\|_2 = \left(|x_1|^2 + \cdots + |x_n|^2\right)^{\frac{1}{2}},$$

and the sup-norm $\|x\|_\infty$, *defined such that,*

$$\|x\|_\infty = \max\{|x_i| \mid 1 \leq i \leq n\}.$$

Some work is required to show the convexity inequality for the Euclidean norm, but this can be found in any standard text.

Note that the Euclidean distance is the distance associated with the Euclidean norm.

One should work out what are the open balls in $\mathbb{R}^2$ for $\| \|_1$ and $\| \|_\infty$. The following proposition is easy to show.

Proposition 1.1. *The following inequalities hold for all $x \in \mathbb{R}^n$ (or $x \in \mathbb{C}^n$):*

$$
\begin{aligned}
\|x\|_\infty &\leq \|x\|_1 \leq n\|x\|_\infty, \\
\|x\|_\infty &\leq \|x\|_2 \leq \sqrt{n}\|x\|_\infty, \\
\|x\|_2 &\leq \|x\|_1 \leq \sqrt{n}\|x\|_2.
\end{aligned}
$$

In a normed vector space, we define a closed ball or an open ball of radius ρ as a closed ball or an open ball of center 0. We may use the notation $B(\rho)$ and $B_0(\rho)$.

We will now define the crucial notions of open sets and closed sets, and of a topological space.

Definition 1.4. *Let E be a metric space with metric d. A subset $U \subseteq E$ is an open set in E if either $U = \emptyset$, or for every $a \in U$, there is some open ball $B_0(a, \rho)$ such that, $B_0(a, \rho) \subseteq U$.*[1] *A subset $F \subseteq E$ is a closed set in E if its complement $E - F$ is open in E.*

The set E itself is open, since for every $a \in E$, every open ball of center a is contained in E.

In $E = \mathbb{R}^n$, given n intervals $[a_i, b_i]$, with $a_i < b_i$, it is easy to show that the open n-cube

$$\{(x_1, \ldots, x_n) \in E \mid a_i < x_i < b_i, 1 \leq i \leq n\}$$

is an open set.

[1] Recall that $\rho > 0$.

In fact, it is possible to find a metric for which such open n-cubes are open balls!

Similarly, we can define the closed n-cube

$$\{(x_1, \ldots, x_n) \in E \mid a_i \leq x_i \leq b_i, \, 1 \leq i \leq n\},$$

which is a closed set.

The open sets satisfy some important properties that lead to the definition of a topological space.

Proposition 1.2. *Given a metric space E with metric d, the family $\mathcal{O}$ of all open sets defined in Definition 1.4 satisfies the following properties:*

(O1) For every finite family $(U_i)_{1\leq i\leq n}$ of sets $U_i \in \mathcal{O}$, we have $U_1 \cap \cdots \cap U_n \in \mathcal{O}$, i.e., $\mathcal{O}$ is closed under finite intersections.

(O2) For every arbitrary family $(U_i)_{i\in I}$ of sets $U_i \in \mathcal{O}$, we have $\bigcup_{i\in I} U_i \in \mathcal{O}$, i.e., $\mathcal{O}$ is closed under arbitrary unions.

(O3) $\emptyset \in \mathcal{O}$, and $E \in \mathcal{O}$, i.e., $\emptyset$ and E belong to $\mathcal{O}$.

Furthermore, for any two distinct points $a \neq b$ in E, there exist two open sets U_a and U_b such that, $a \in U_a$, $b \in U_b$, and $U_a \cap U_b = \emptyset$.

The above proposition leads to the very general concept of a topological space.

One should be careful that, in general, the family of open sets is not closed under infinite intersections.

For example, in $\mathbb{R}$ under the metric $|x-y|$, letting $U_n = (-1/n, +1/n)$, each U_n is open, but $\bigcap_n U_n = \{0\}$, which is not open.

1.2 Topological Spaces

Definition 1.5. *Given a set E, a topology on E (or a topological structure on E), is defined as a family $\mathcal{O}$ of subsets of E called open sets, and satisfying the following three properties:*

(1) For every finite family $(U_i)_{1\leq i\leq n}$ of sets $U_i \in \mathcal{O}$, we have $U_1 \cap \cdots \cap U_n \in \mathcal{O}$, i.e., $\mathcal{O}$ is closed under finite intersections.

(2) For every arbitrary family $(U_i)_{i\in I}$ of sets $U_i \in \mathcal{O}$, we have $\bigcup_{i\in I} U_i \in \mathcal{O}$, i.e., $\mathcal{O}$ is closed under arbitrary unions.

(3) $\emptyset \in \mathcal{O}$, and $E \in \mathcal{O}$, i.e., $\emptyset$ and E belong to $\mathcal{O}$.

A set E together with a topology $\mathcal{O}$ on E is called a topological space. Given a topological space $(E, \mathcal{O})$, a subset F of E is a closed set if $F = E - U$ for some open set $U \in \mathcal{O}$, i.e., F is the complement of some open set.

It is possible that an open set is also a closed set. For example, $\emptyset$ and E are both open and closed. When a topological space contains a proper nonempty subset U which is both open and closed, the space E is said to be disconnected.

A topological space $(E, \mathcal{O})$ is said to satisfy the Hausdorff separation axiom (or T_2-separation axiom) if for any two distinct points $a \neq b$ in E, there exist two open sets U_a and U_b such that, $a \in U_a$, $b \in U_b$, and $U_a \cap U_b = \emptyset$. When the T_2-separation axiom is satisfied, we also say that $(E, \mathcal{O})$ is a Hausdorff space.

Sometimes, it is more convenient to define a topology in terms of its *closed sets*.

As shown by Proposition 1.2, any metric space is a topological Hausdorff space, the family of open sets being in fact the family of arbitrary unions of open balls.

Similarly, any normed vector space is a topological Hausdorff space, the family of open sets being the family of arbitrary unions of open balls. The topology $\mathcal{O}$ consisting of all subsets of E is called the *discrete topology*.

Remark: Most (if not all) spaces used in analysis are Hausdorff spaces. Intuitively, the Hausdorff separation axiom says that there are enough "small" open sets.

Without this axiom, some counter-intuitive behaviors may arise. For example, a sequence may have more than one limit point (or a compact set may not be closed).

Nevertheless, non-Hausdorff topological spaces arise naturally in algebraic geometry. In the *Zariski topology*, the closed sets are the zero loci of sets of algebraic equations. But even there, some substitute for separation is used.

One of the reasons why topological spaces are important is that the definition of a topology only involves a certain family $\mathcal{O}$ of sets, and not **how** such family is generated from a metric or a norm.

For example, different metrics or different norms can define the same family of open sets. Many topological properties only depend on the family $\mathcal{O}$ and not on the specific metric or norm.

But the fact that a topology is definable from a metric or a norm is important, because it usually implies nice properties of a space.

All our examples will be spaces whose topology is defined by a metric or a norm.

By taking complements, we can state properties of the closed sets dual to those of Definition 1.5. Thus, $\emptyset$ and E are closed sets, and the closed sets are closed under finite unions and arbitrary intersections.

It is also worth noting that the Hausdorff separation axiom implies that for every $a \in E$, the set $\{a\}$ is closed.

Given a topological space $(E, \mathcal{O})$, given any subset A of E, since $E \in \mathcal{O}$ and E is a closed set, the family

$$\mathcal{C}_A = \{F \mid A \subseteq F,\ F \text{ a closed set}\}$$

of closed sets containing A is nonempty, and since any arbitrary intersection of closed sets is a closed set, the intersection $\bigcap \mathcal{C}_A$ of the sets in the family $\mathcal{C}_A$ is the smallest closed set containing A.

By a similar reasoning, the union of all the open subsets contained in A is the largest open set contained in A.

Definition 1.6. *Given a topological space* $(E, \mathcal{O})$*, given any subset* A *of* E*, the smallest closed set containing* A *is denoted by* $\overline{A}$*, and is called the closure, or adherence of* A*. A subset* A *of* E *is dense in* E *if* $\overline{A} = E$*. The largest open set contained in* A *is denoted by* $\overset{\circ}{A}$*, and is called the interior of* A*. The set* $\mathrm{Fr}\, A = \overline{A} \cap \overline{E - A}$ *is called the boundary (or frontier) of* A*. We also denote the boundary of* A *by* ∂A*.*

Remark: The notation $\overline{A}$ for the closure of a subset A of E is somewhat unfortunate, since $\overline{A}$ is often used to denote the set complement of A in E.

Still, we prefer it to more cumbersome notations such as $\mathrm{clo}(A)$, and we denote the complement of A in E by $E - A$ (or sometimes, A^c).

By definition, it is clear that a subset A of E is closed iff $A = \overline{A}$. The set $\mathbb{Q}$ of rationals is dense in $\mathbb{R}$.

It is easily shown that $\overline{A} = \overset{\circ}{A} \cup \partial A$ and $\overset{\circ}{A} \cap \partial A = \emptyset$. Another useful characterization of $\overline{A}$ is given by the following proposition.

Proposition 1.3. *Given a topological space* $(E, \mathcal{O})$*, given any subset* A *of* E*, the closure* $\overline{A}$ *of* A *is the set of all points* $x \in E$ *such that for every open set* U *containing* x*, then* $U \cap A \neq \emptyset$*.*

Often, it is necessary to consider a subset A of a topological space E, and to view the subset A as a topological space. The following proposition shows how to define a topology on a subset.

Proposition 1.4. *Given a topological space* $(E, \mathcal{O})$*, given any subset* A *of* E*, let*

$$\mathcal{U} = \{U \cap A \mid U \in \mathcal{O}\}$$

be the family of all subsets of A *obtained as the intersection of any open set in* $\mathcal{O}$ *with* A*. The following properties hold.*

(1) The space $(A, \mathcal{U})$ *is a topological space.*

(2) If E *is a metric space with metric* d*, then the restriction* $d_A \colon A \times A \to \mathbb{R}_+$ *of the metric* d *to* A *defines a metric space. Furthermore, the topology induced by the metric* d_A *agrees with the topology defined by* $\mathcal{U}$*, as above.*

Proposition 1.4 suggests the following definition.

Definition 1.7. *Given a topological space* $(E, \mathcal{O})$*, given any subset* A *of* E*, the subspace topology on* A *induced by* $\mathcal{O}$ *is the family* $\mathcal{U}$ *of open sets defined such that*

$$\mathcal{U} = \{U \cap A \mid U \in \mathcal{O}\}$$

is the family of all subsets of A *obtained as the intersection of any open set in* $\mathcal{O}$ *with* A*. We say that* $(A, \mathcal{U})$ *has the subspace topology. If* (E, d) *is a metric space, the restriction* $d_A \colon A \times A \to \mathbb{R}_+$ *of the metric* d *to* A *is called the subspace metric.*

For example, if $E = \mathbb{R}^n$ and d is the Euclidean metric, we obtain the subspace topology on the closed n-cube

$$\{(x_1, \ldots, x_n) \in E \mid a_i \leq x_i \leq b_i, 1 \leq i \leq n\}.$$

One should realize that every open set $U \in \mathcal{O}$ which is entirely contained in A is also in the family $\mathcal{U}$, but $\mathcal{U}$ may contain open sets that are not in $\mathcal{O}$.

For example, if $E = \mathbb{R}$ with $|x - y|$, and $A = [a, b]$, then sets of the form $[a, c)$, with $a < c < b$ belong to $\mathcal{U}$, but they are not open sets for $\mathbb{R}$ under $|x - y|$. However, there is agreement in the following situation.

Proposition 1.5. *Given a topological space $(E, \mathcal{O})$, given any subset A of E, if $\mathcal{U}$ is the subspace topology, then the following properties hold.*

(1) If A is an open set $A \in \mathcal{O}$, then every open set $U \in \mathcal{U}$ is an open set $U \in \mathcal{O}$.

(2) If A is a closed set in E, then every closed set w.r.t. the subspace topology is a closed set w.r.t. $\mathcal{O}$.

The concept of product topology is also useful. We have the following proposition.

Proposition 1.6. *Given n topological spaces $(E_i, \mathcal{O}_i)$, let $\mathcal{B}$ be the family of subsets of $E_1 \times \cdots \times E_n$ defined as follows:*

$$\mathcal{B} = \{U_1 \times \cdots \times U_n \mid U_i \in \mathcal{O}_i,\ 1 \leq i \leq n\},$$

and let $\mathcal{P}$ be the family consisting of arbitrary unions of sets in $\mathcal{B}$, including $\emptyset$. Then, $\mathcal{P}$ is a topology on $E_1 \times \cdots \times E_n$.

Definition 1.8. *Given n topological spaces $(E_i, \mathcal{O}_i)$, the product topology on $E_1 \times \cdots \times E_n$ is the family $\mathcal{P}$ of subsets of $E_1 \times \cdots \times E_n$ defined as follows: if*

$$\mathcal{B} = \{U_1 \times \cdots \times U_n \mid U_i \in \mathcal{O}_i,\ 1 \leq i \leq n\},$$

then $\mathcal{P}$ is the family consisting of arbitrary unions of sets in $\mathcal{B}$, including $\emptyset$.

If each $(E_i, \| \|_i)$ is a normed vector space, there are three natural norms that can be defined on $E_1 \times \cdots \times E_n$:

$$\|(x_1, \ldots, x_n)\|_1 = \|x_1\|_1 + \cdots + \|x_n\|_n,$$
$$\|(x_1, \ldots, x_n)\|_2 = \left(\|x_1\|_1^2 + \cdots + \|x_n\|_n^2\right)^{\frac{1}{2}},$$
$$\|(x_1, \ldots, x_n)\|_\infty = \max\left\{\|x_1\|_1, \ldots, \|x_n\|_n\right\}.$$

It is easy to show that they all define the same topology, which is the product topology.

It can also be verified that when $E_i = \mathbb{R}$, with the standard topology induced by $|x - y|$, the topology product on $\mathbb{R}^n$ is the standard topology induced by the Euclidean norm.

Definition 1.9. *Two metrics d and d' on a space E are equivalent if they induce the same topology $\mathcal{O}$ on E (i.e., they define the same family $\mathcal{O}$ of open sets). Similarly, two norms $\| \|$ and $\| \|'$ on a space E are equivalent if they induce the same topology $\mathcal{O}$ on E.*

Remark: Given a topological space $(E, \mathcal{O})$, it is often useful, as in Proposition 1.6, to define the topology $\mathcal{O}$ in terms of a subfamily $\mathcal{B}$ of subsets of E.

We say that a family $\mathcal{B}$ of subsets of E is a *basis for the topology $\mathcal{O}$*, if $\mathcal{B}$ is a subset of $\mathcal{O}$, and if every open set U in $\mathcal{O}$ can be obtained as some union (possibly infinite) of sets in $\mathcal{B}$ (agreeing that the empty union is the empty set).

A *subbasis for $\mathcal{O}$* is a family $\mathcal{S}$ of subsets of E, such that the family $\mathcal{B}$ of all finite intersections of sets in $\mathcal{S}$ (including E itself, in case of the empty intersection) is a basis of $\mathcal{O}$.

We now consider the fundamental property of continuity.

1.3 Continuous Functions, Limits

Definition 1.10. *Let $(E, \mathcal{O}_E)$ and $(F, \mathcal{O}_F)$ be topological spaces, and let $f \colon E \to F$ be a function. For every $a \in E$, we say that f is continuous at a, if for every open set $V \in \mathcal{O}_F$ containing $f(a)$, there is some open set $U \in \mathcal{O}_E$ containing a, such that, $f(U) \subseteq V$. We say that f is continuous if it is continuous at every $a \in E$.*

Define a *neighborhood of $a \in E$* as any subset N of E containing some open set $O \in \mathcal{O}$ such that $a \in O$.

Now, if f is continuous at a and N is any neighborhood of $f(a)$, there is some open set $V \subseteq N$ containing $f(a)$, and since f is continuous at a, there is some open set U containing a, such that $f(U) \subseteq V$.

Since $V \subseteq N$, the open set U is a subset of $f^{-1}(N)$ containing a, and $f^{-1}(N)$ is a neighborhood of a.

Conversely, if $f^{-1}(N)$ is a neighborhood of a whenever N is any neighborhood of $f(a)$, it is immediate that f is continuous at a.

It is easy to see that Definition 1.10 is equivalent to the following statements.

Proposition 1.7. *Let $(E, \mathcal{O}_E)$ and $(F, \mathcal{O}_F)$ be topological spaces, and let $f : E \to F$ be a function. For every $a \in E$, the function f is continuous at $a \in E$ iff for every neighborhood N of $f(a) \in F$, then $f^{-1}(N)$ is a neighborhood of a. The function f is continuous on E iff $f^{-1}(V)$ is an open set in $\mathcal{O}_E$ for every open set $V \in \mathcal{O}_F$.*

If E and F are metric spaces defined by metrics d_E and d_F, we can show easily that *f is continuous at a* iff

for every $\epsilon > 0$, there is some $\eta > 0$, such that, for every $x \in E$,

$$\text{if } d_E(a, x) \leq \eta, \text{ then } d_F(f(a), f(x)) \leq \epsilon.$$

Similarly, if E and F are normed vector spaces defined by norms $\| \|_E$ and $\| \|_F$, we can show easily that *f is continuous at a* iff

for every $\epsilon > 0$, there is some $\eta > 0$, such that, for every $x \in E$,

$$\text{if } \|x - a\|_E \leq \eta, \text{ then } \|f(x) - f(a)\|_F \leq \epsilon.$$

It is worth noting that continuity is a topological notion, in the sense that equivalent metrics (or equivalent norms) define exactly the same notion of continuity.

If $(E, \mathcal{O}_E)$ and $(F, \mathcal{O}_F)$ are topological spaces, and $f\colon E \to F$ is a function, for every nonempty subset $A \subseteq E$ of E, we say that *f is continuous on A* if the restriction of f to A is continuous with respect to $(A, \mathcal{U})$ and $(F, \mathcal{O}_F)$, where $\mathcal{U}$ is the subspace topology induced by $\mathcal{O}_E$ on A.

Given a product $E_1 \times \cdots \times E_n$ of topological spaces, as usual, we let $\pi_i\colon E_1 \times \cdots \times E_n \to E_i$ be the projection function such that, $\pi_i(x_1, \ldots, x_n) = x_i$. It is immediately verified that each π_i is continuous.

Given a topological space $(E, \mathcal{O})$, we say that a point $a \in E$ is *isolated* if $\{a\}$ is an open set in $\mathcal{O}$.

Then, if $(E, \mathcal{O}_E)$ and $(F, \mathcal{O}_F)$ are topological spaces, any function $f\colon E \to F$ is continuous at every isolated point $a \in E$. In the discrete topology, every point is isolated.

In a nontrivial normed vector space $(E, \| \, \|)$ (with $E \neq \{0\}$), no point is isolated.

The following proposition is easily shown.

Proposition 1.8. *Given topological spaces* $(E, \mathcal{O}_E)$, $(F, \mathcal{O}_F)$, *and* $(G, \mathcal{O}_G)$, *and two functions* $f\colon E \to F$ *and* $g\colon F \to G$, *if* f *is continuous at* $a \in E$ *and* g *is continuous at* $f(a) \in F$, *then* $g \circ f\colon E \to G$ *is continuous at* $a \in E$. *Given* n *topological spaces* $(F_i, \mathcal{O}_i)$, *for every function* $f\colon E \to F_1 \times \cdots \times F_n$, *then* f *is continuous at* $a \in E$ *iff every* $f_i\colon E \to F_i$ *is continuous at* a, *where* $f_i = \pi_i \circ f$.

One can also show that in a metric space (E, d), the metric $d\colon E \times E \to \mathbb{R}$ is continuous, where $E \times E$ has the product topology, and that for a normed vector space $(E, \| \, \|)$, the norm $\| \, \|\colon E \to \mathbb{R}$ is continuous.

Given a function $f\colon E_1 \times \cdots \times E_n \to F$, we can fix $n-1$ of the arguments, say $a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n$, and view f as a function of the remaining argument,

$$x_i \mapsto f(a_1, \ldots, a_{i-1}, x_i, a_{i+1}, \ldots, a_n),$$

where $x_i \in E_i$. If f is continuous, it is clear that each f_i is continuous.

One should be careful that the converse is false! For example, consider the function $f\colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, defined such that,

$$f(x, y) = \frac{xy}{x^2 + y^2} \quad \text{if } (x, y) \neq (0, 0), \quad \text{and} \quad f(0, 0) = 0.$$

The function f is continuous on $\mathbb{R} \times \mathbb{R} - \{(0, 0)\}$, but on the line $y = mx$, with $m \neq 0$, we have $f(x, y) = \frac{m}{1+m^2} \neq 0$, and thus, on this line, $f(x, y)$ does not approach 0 when (x, y) approaches $(0, 0)$.

The following proposition is useful for showing that real-valued functions are continuous.

Proposition 1.9. *If E is a topological space, and $(\mathbb{R}, |x - y|)$ the reals under the standard topology, for any two functions $f\colon E \to \mathbb{R}$ and $g\colon E \to \mathbb{R}$, for any $a \in E$, for any $\lambda \in \mathbb{R}$, if f and g are continuous at a, then $f + g$, λf, $f \cdot g$, are continuous at a, and f/g is continuous at a if $g(a) \neq 0$.*

Using Proposition 1.9, we can show easily that every real polynomial function is continuous.

The notion of isomorphism of topological spaces is defined as follows.

Definition 1.11. *Let $(E, \mathcal{O}_E)$ and $(F, \mathcal{O}_F)$ be topological spaces, and let $f\colon E \to F$ be a function. We say that f is a homeomorphism between E and F if f is bijective, and both $f\colon E \to F$ and $f^{-1}\colon F \to E$ are continuous.*

One should be careful that a bijective continuous function $f\colon E \to F$ is not necessarily an homeomorphism.

For example, if $E = \mathbb{R}$ with the discrete topology, and $F = \mathbb{R}$ with the standard topology, the identity is not a homeomorphism.

Another interesting example involving a parametric curve is given below. Let $L\colon \mathbb{R} \to \mathbb{R}^2$ be the function, defined such that,

$$L_1(t) = \frac{t(1+t^2)}{1+t^4},$$
$$L_2(t) = \frac{t(1-t^2)}{1+t^4}.$$

If we think of $(x(t), y(t)) = (L_1(t), L_2(t))$ as a geometric point in $\mathbb{R}^2$, the set of points $(x(t), y(t))$ obtained by letting t vary in $\mathbb{R}$ from $-\infty$ to $+\infty$, defines a curve having the shape of a "figure eight", with self-intersection at the origin, called the "lemniscate of Bernoulli."

The map L is continuous, and in fact bijective, but its inverse L^{-1} is not continuous.

Indeed, when we approach the origin on the branch of the curve in the upper left quadrant (i.e., points such that, $x \leq 0, y \geq 0$), then t goes to $-\infty$, and when we approach the origin on the branch of the curve in the lower right quadrant (i.e., points such that, $x \geq 0$, $y \leq 0$), then t goes to $+\infty$.

We also review the concept of limit of a sequence. Given any set E, a *sequence* is any function $x\colon \mathbb{N} \to E$, usually denoted by $(x_n)_{n\in\mathbb{N}}$, or $(x_n)_{n\geq 0}$, or even by (x_n).

Definition 1.12. *Given a topological space* $(E, \mathcal{O})$*, we say that a sequence* $(x_n)_{n\in\mathbb{N}}$ *converges to some* $a \in E$ *if for every open set* U *containing* a*, there is some* $n_0 \geq 0$*, such that,* $x_n \in U$*, for all* $n \geq n_0$*. We also say that* a *is a limit of* $(x_n)_{n\in\mathbb{N}}$*.*

When E is a metric space with metric d, it is easy to show that this is equivalent to the fact that,

for every $\epsilon > 0$, there is some $n_0 \geq 0$, such that,
$d(x_n, a) \leq \epsilon$, for all $n \geq n_0$.

When E is a normed vector space with norm $\| \: \|$, it is easy to show that this is equivalent to the fact that,

for every $\epsilon > 0$, there is some $n_0 \geq 0$, such that,
$\|x_n - a\| \leq \epsilon$, for all $n \geq n_0$.

The following proposition shows the importance of the Hausdorff separation axiom.

Proposition 1.10. *Given a topological space* $(E, \mathcal{O})$*, if the Hausdorff separation axiom holds, then every sequence has at most one limit.*

It is worth noting that the notion of limit is topological, in the sense that a sequence converge to a limit b iff it converges to the same limit b in any equivalent metric (and similarly for equivalent norms).

We still need one more concept of limit for functions.

Definition 1.13. *Let $(E, \mathcal{O}_E)$ and $(F, \mathcal{O}_F)$ be topological spaces, let A be some nonempty subset of E, and let $f\colon A \to F$ be a function. For any $a \in \overline{A}$ and any $b \in F$, we say that $f(x)$ approaches b as x approaches a with values in A if for every open set $V \in \mathcal{O}_F$ containing b, there is some open set $U \in \mathcal{O}_E$ containing a, such that, $f(U \cap A) \subseteq V$. This is denoted by*

$$\lim_{x \to a, x \in A} f(x) = b.$$

First, note that by Proposition 1.3, since $a \in \overline{A}$, for every open set U containing a, we have $U \cap A \neq \emptyset$, and the definition is nontrivial. Also, even if $a \in A$, the value $f(a)$ of f at a plays no role in this definition.

When E and F are metric space with metrics d_E and d_F, it can be shown easily that the definition can be stated as follows:

For every $\epsilon > 0$, there is some $\eta > 0$, such that, for every $x \in A$,

$$\text{if } d_E(x, a) \leq \eta, \text{ then } d_F(f(x), b) \leq \epsilon.$$

When E and F are normed vector spaces with norms $\| \|_E$ and $\| \|_F$, it can be shown easily that the definition can be stated as follows:

For every $\epsilon > 0$, there is some $\eta > 0$, such that, for every $x \in A$,

$$\text{if } \|x - a\|_E \leq \eta, \text{ then } \|f(x) - b\|_F \leq \epsilon.$$

We have the following result relating continuity at a point and the previous notion.

Proposition 1.11. *Let $(E, \mathcal{O}_E)$ and $(F, \mathcal{O}_F)$ be two topological spaces, and let $f: E \to F$ be a function. For any $a \in E$, the function f is continuous at a iff $f(x)$ approaches $f(a)$ when x approaches a (with values in E).*

Another important proposition relating the notion of convergence of a sequence to continuity, is stated without proof.

Proposition 1.12. *Let $(E, \mathcal{O}_E)$ and $(F, \mathcal{O}_F)$ be two topological spaces, and let $f: E \to F$ be a function.*

(1) If f is continuous, then for every sequence $(x_n)_{n\in\mathbb{N}}$ in E, if (x_n) converges to a, then $(f(x_n))$ converges to $f(a)$.

(2) If E is a metric space, and $(f(x_n))$ converges to $f(a)$ whenever (x_n) converges to a, for every sequence $(x_n)_{n\in\mathbb{N}}$ in E, then f is continuous.

A special case of Definition 1.13 will be used when E and F are (nontrivial) normed vector spaces with norms $\| \, \|_1$ and $\| \, \|_2$.

Let U be any nonempty open subset of E. We showed earlier that E has no isolated points and that every set $\{v\}$ is closed, for every $v \in E$.

Since E is nontrivial, for every $v \in U$, there is a nontrivial open ball contained in U (an open ball not reduced to its center).

Then, for every $v \in U$, $A = U - \{v\}$ is open and nonempty, and clearly, $v \in \overline{A}$.

For any $v \in U$, if $f(x)$ approaches b when x approaches v with values in $A = U - \{v\}$, we say that *$f(x)$ approaches b when x approaches v with values $\neq v$ in U.*

This is denoted by

$$\lim_{x \to v, x \in U, x \neq v} f(x) = b.$$

Remark: Variations of the above case show up in the following case: $E = \mathbb{R}$, and F is some arbitrary topological space.

Let A be some nonempty subset of $\mathbb{R}$, and let $f\colon A \to F$ be some function. For any $a \in A$, we say that *f is continuous on the right at a* if

$$\lim_{x \to a, x \in A \cap [a, +\infty)} f(x) = f(a).$$

We can define *continuity on the left* at a in a similar fashion.

Let us consider another variation. Let A be some nonempty subset of $\mathbb{R}$, and let $f\colon A \to F$ be some function.

For any $a \in A$, we say that *f has a discontinuity of the first kind at a* if

$$\lim_{x \to a, x \in A \cap (-\infty, a)} f(x) = f(a_-)$$

and

$$\lim_{x \to a, x \in A \cap (a, +\infty)} f(x) = f(a_+)$$

both exist, and either $f(a_-) \neq f(a)$, or $f(a_+) \neq f(a)$.

Note that it is possible that $f(a_-) = f(a_+)$, but f is still discontinuous at a if this common value differs from $f(a)$.

Functions defined on a nonempty subset of $\mathbb{R}$, and that are continuous, except for some points of discontinuity of the first kind, play an important role in analysis.

In a metric space there is another important notion of continuity, namely uniform continuity.

Definition 1.14. Given two metric spaces, (E, d_E) and (F, d_F), a function $f\colon E \to F$ is *uniformly continuous* if for every $\epsilon > 0$, there is some $\eta > 0$, such that for all $a, b \in E$,

$$\text{if} \quad d_E(a, b) \leq \eta \quad \text{then} \quad d_F(f(a), f(b)) \leq \epsilon.$$

See Figures 1.1 and 1.2.

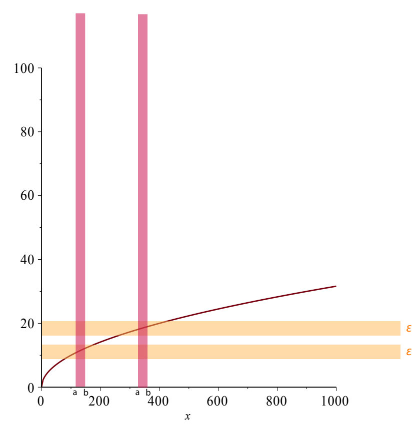

Figure 1.1: The real valued function $f(x) = \sqrt{x}$ is uniformly continuous over $(0, \infty)$. Fix ϵ. If the x values lie within the rose colored η strip, the y values always lie within the peach ϵ strip.

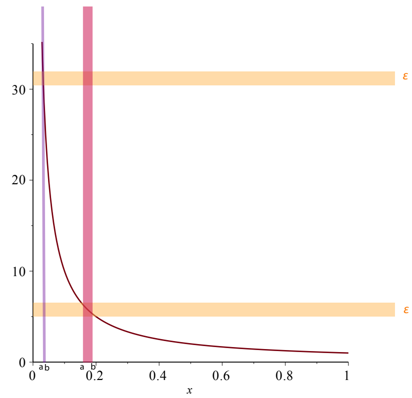

Figure 1.2: The real valued function $f(x) = 1/x$ is not uniformly continuous over $(0, \infty)$. Fix ϵ. In order for the y values to lie within the peach epsilon strip, the widths of the eta strips decrease as $x \to 0$.

It is easily shown that the metric on a metric space is uniformly continuous, and the norm on a normed metric space is uniformly continuous.

Before considering differentials, we need to look at the continuity of linear maps.

1.4 Continuous Linear and Multilinear Maps

If E and F are normed vector spaces, we first characterize when a linear map $f\colon E \to F$ is continuous.

Proposition 1.13. *Given two normed vector spaces E and F, for any linear map $f\colon E \to F$, the following conditions are equivalent:*

(1) The function f is continuous at 0.

(2) There is a constant $k \geq 0$ such that,

$$\|f(u)\| \leq k, \text{ for every } u \in E \text{ such that } \|u\| \leq 1.$$

(3) There is a constant $k \geq 0$ such that,

$$\|f(u)\| \leq k\|u\|, \text{ for every } u \in E.$$

(4) The function f is continuous at every point of E.

Among other things, Proposition 1.13 shows that a linear map is continuous iff the image of the unit (closed) ball is bounded.

If E and F are normed vector spaces, the set of all continuous linear maps $f\colon E \to F$ is denoted by $\mathcal{L}(E; F)$.

Using Proposition 1.13, we can define a norm on $\mathcal{L}(E; F)$ which makes it into a normed vector space.

Definition 1.15. Given two normed vector spaces E and F, for every continuous linear map $f\colon E \to F$, we define the *(operator) norm $\|f\|$ of f* as

$$\begin{aligned}\|f\| &= \inf \{k \geq 0 \mid \|f(x)\| \leq k\|x\|, \text{ for all } x \in E\} \\ &= \sup \{\|f(x)\| \mid \|x\| \leq 1\} \\ &= \sup \{\|f(x)\| \mid \|x\| = 1\}.\end{aligned}$$

From Definition 1.15, for every continuous linear map $f \in \mathcal{L}(E; F)$, we have

$$\|f(x)\| \leq \|f\| \|x\|,$$

for every $x \in E$.

The above implies that a continuous linear map is actually *uniformly continuous*.

It is easy to verify that $\mathcal{L}(E; F)$ is a normed vector space under the norm of Definition 1.15.

Furthermore, if E, F, G, are normed vector spaces, and $f \colon E \to F$ and $g \colon F \to G$ are continuous linear maps, we have

$$\|g \circ f\| \leq \|g\| \|f\|.$$

We can now show that when $E = \mathbb{R}^n$ or $E = \mathbb{C}^n$, with any of the norms $\| \|_1$, $\| \|_2$, or $\| \|_\infty$, then every linear map $f\colon E \to F$ is continuous.

Proposition 1.14. *If $E = \mathbb{R}^n$ or $E = \mathbb{C}^n$, with any of the norms $\| \|_1$, $\| \|_2$, or $\| \|_\infty$, and F is any normed vector space, then every linear map $f\colon E \to F$ is continuous.*

Actually, it can be shown that if E is a vector space of finite dimension, then any two norms define the same topology. We say that any two norms are equivalent.

Proposition 1.15. *If E is a vector space of finite dimension (over $\mathbb{R}$ or $\mathbb{C}$), then all norms are equivalent (define the same topology). Furthermore, for any normed vector space F, every linear map $f\colon E \to F$ is continuous.*

If E is a normed vector space of infinite dimension, a linear map $f\colon E \to F$ may not be continuous.

As an example, let E be the infinite vector space of all polynomials over $\mathbb{R}$. Let

$$\|P(X)\| = \max_{0 \leq x \leq 1} |P(x)|.$$

We leave as an exercise to show that this is indeed a norm.

Let $F = \mathbb{R}$, and let $f\colon E \to F$ be the map defined such that, $f(P(X)) = P(3)$. It is clear that f is linear.

Consider the sequence of polynomials

$$P_n(X) = \left(\frac{X}{2}\right)^n.$$

It is clear that $\|P_n\| = \left(\frac{1}{2}\right)^n$, and thus, the sequence P_n has the null polynomial as a limit.

However, we have

$$f(P_n(X)) = P_n(3) = \left(\frac{3}{2}\right)^n,$$

and the sequence $f(P_n(X))$ diverges to $+\infty$. Consequently, in view of Proposition 1.12 (1), f is not continuous.

We now consider the continuity of multilinear maps. We treat explicitly bilinear maps, the general case being a straightforward extension.

Proposition 1.16. *Given normed vector spaces E, F and G, for any bilinear map $f\colon E \times F \to G$, the following conditions are equivalent:*

(1) The function f is continuous at $\langle 0, 0\rangle$.

(2) There is a constant $k \geq 0$ such that,

$$\|f(u,v)\| \leq k, \ \text{for all } u \in E, v \in F \text{ such that } \|u\|, \|v\| \leq 1.$$

(3) There is a constant $k \geq 0$ such that,

$$\|f(u,v)\| \leq k\|u\|\|v\|, \ \text{for all } u \in E, v \in F.$$

(4) The function f is continuous at every point of $E \times F$.

If E, F, and G, are normed vector spaces, we denote the set of all continuous bilinear maps $f\colon E \times F \to G$ by $\mathcal{L}_2(E, F; G)$.

Using Proposition 1.16, we can define a norm on $\mathcal{L}_2(E, F; G)$ which makes it into a normed vector space.

Definition 1.16. Given normed vector spaces E, F, and G, for every continuous bilinear map $f \colon E \times F \to G$, we define the *(operator) norm $\|f\|$ of f* as

$$\begin{aligned} \|f\| &= \inf\{k \geq 0 \mid \|f(x,y)\| \leq k\|x\|\|y\|, \;\; x \in E, y \in F\} \\ &= \sup\{\|f(x,y)\| \mid \|x\|, \|y\| \leq 1\} \\ &= \sup\{\|f(x,y)\| \mid \|x\| = \|y\| = 1\}. \end{aligned}$$

From Definition 1.16, for every continuous bilinear map $f \in \mathcal{L}_2(E, F; G)$, we have

$$\|f(x,y)\| \leq \|f\|\|x\|\|y\|,$$

for all $x \in E, y \in F$.

It is easy to verify that $\mathcal{L}_2(E, F; G)$ is a normed vector space under the norm of Definition 1.16.

In contrast to continuous linear maps, which must be uniformly continuous, nonzero continuous bilinear maps are *not* uniformly continuous; see our book, Section 2.6, Vol. II.

Given a bilinear map $f\colon E \times F \to G$, for every $u \in E$, we obtain a linear map denoted $fu\colon F \to G$, defined such that, $fu(v) = f(u, v)$.

Furthermore, since

$$\|f(x, y)\| \leq \|f\|\|x\|\|y\|,$$

it is clear that fu is continuous.

We can then consider the map $\varphi\colon E \to \mathcal{L}(F; G)$, defined such that, $\varphi(u) = fu$, for any $u \in E$, or equivalently, such that,

$$\varphi(u)(v) = f(u, v).$$

Actually, it is easy to show that φ is linear and continuous, and that $\|\varphi\| = \|f\|$.

Thus, $f \mapsto \varphi$ defines a map from $\mathcal{L}_2(E, F; G)$ to $\mathcal{L}(E; \mathcal{L}(F; G))$. We can also go back from $\mathcal{L}(E; \mathcal{L}(F; G))$ to $\mathcal{L}_2(E, F; G)$.

We summarize all this in the following proposition.

Proposition 1.17. *Let E, F, G be normed vector spaces. The map $f \mapsto \varphi$, from $\mathcal{L}_2(E, F; G)$ to $\mathcal{L}(E; \mathcal{L}(F; G))$, defined such that, for every $f \in \mathcal{L}_2(E, F; G)$,*

$$\varphi(u)(v) = f(u, v),$$

is an isomorphism of vector spaces, and furthermore, $\|\varphi\| = \|f\|$.

As a corollary of Proposition 1.17, we get the following proposition which will be useful when we define second-order derivatives.

Proposition 1.18. *Let E, F be normed vector spaces. The map app from $\mathcal{L}(E; F) \times E$ to F, defined such that, for every $f \in \mathcal{L}(E; F)$, for every $u \in E$,*

$$app(f, u) = f(u),$$

is a continuous bilinear map.

Remark: If E and F are nontrivial, it can be shown that $\|\mathrm{app}\| = 1$. It can also be shown that composition

$$\circ \colon \mathcal{L}(E;F) \times \mathcal{L}(F;G) \to \mathcal{L}(E;G),$$

is bilinear and continuous.

The above propositions and definition generalize to arbitrary n-multilinear maps, with $n \geq 2$.

Proposition 1.16 extends in the obvious way to any n-multilinear map $f \colon E_1 \times \cdots \times E_n \to F$, but condition (3) becomes:

There is a constant $k \geq 0$ such that,

$$\|f(u_1, \ldots, u_n)\| \leq k\|u_1\| \cdots \|u_n\|,$$

for all $u_1 \in E_1, \ldots, u_n \in E_n$

Definition 1.16 also extends easily to

$$\begin{aligned}\|f\| &= \min\Big\{k \geq 0 \mid \|f(x_1, \ldots, x_n)\| \leq k\|x_1\| \cdots \|x_n\|, \\ &\qquad\qquad \text{for all } x_i \in E_i, 1 \leq i \leq n\Big\} \\ &= \sup\left\{\|f(x_1, \ldots, x_n)\| \mid \|x_1\|, \ldots, \|x_n\| \leq 1\right\} \\ &= \sup\left\{\|f(x_1, \ldots, x_n)\| \mid \|x_1\| = \cdots = \|x_n\| = 1\right\}.\end{aligned}$$

Proposition 1.17 is also easily extended, and we get an isomorphism between continuous n-multilinear maps in $\mathcal{L}_n(E_1, \ldots, E_n; F)$, and continuous linear maps in

$$\mathcal{L}(E_1; \mathcal{L}(E_2; \ldots; \mathcal{L}(E_n; F)))$$

An obvious extension of Proposition 1.18 also holds.

For the sake of completeness, we include the definition of Cauchy sequences.

Definition 1.17. Given a metric space (E, d), a sequence $(x_n)_{n\in\mathbb{N}}$ in E is a *Cauchy sequence* if the following condition holds: For every $\epsilon > 0$, there is some $p \geq 0$, such that, for all $m, n \geq p$, then $d(x_m, x_n) \leq \epsilon$.

If every Cauchy sequence in (E, d) converges, we say that (E, d) is a *complete metric space*. A normed vector space $(E, \| \, \|)$ over $\mathbb{R}$ (or $\mathbb{C}$) which is a complete metric space for the distance $\|v - u\|$, is called a *Banach space*.

The standard example of a complete metric space is the set $\mathbb{R}$ of real numbers.

As a matter of fact, the set $\mathbb{R}$ can be defined as the "completion" of the set $\mathbb{Q}$ of rationals.

The spaces $\mathbb{R}^n$ and $\mathbb{C}^n$ under their standard topology are complete metric spaces.

It can be shown that every normed vector space of finite dimension is a Banach space (is complete).

It can also be shown that if E and F are normed vector spaces, and F is a Banach space, then $\mathcal{L}(E; F)$ is a Banach space.

If E, F and G are normed vector spaces, and G is a Banach space, then $\mathcal{L}_2(E, F; G)$ is a Banach space.

1.5 Futher Readings

A thorough treatment of general topology can be found in Munkres [7, 6], Dixmier [3], Lang [5], Schwartz [10, 9], Bredon [1] and the classic, Seifert and Threlfall [11].