Norms, Metrics and Topologies

Lecture Notes

Rashid M. Talha

School of Natural Sciences, NUST

(Date: December 13, 2022)

Notes.

First delivered at NUST during the fall semester, 2022–23. Currently incomplete.

1. Review, terminology, and notation

BASICS OF SETS. SOME NOTATION.

BASICS OF FUNCTIONS. THE INVERSE-IMAGE (PRE-IMAGE). THE EXISTENCE AND UNIQUENESS OF THE INVERSE. SOME NOTATION.

BASICS OF INEQUALITIES.

RECAP OF THE DEFINITION OF SEQUENCES AND THEIR LIMITS. DEFINITION OF A CAUCHY SEQUENCE.

2. Normed vector spaces

Definition 2.1 (Norm)**.** Let V be a vector space over $\mathbb{R}$. A norm on V is a function $\|\cdot\| : V \to \mathbb{R}$ with the following properties

(i) $\|v\| = 0 \iff v = 0$, for $v \in V$, (definiteness)

(ii) $\|\lambda v\| = |\lambda|\|v\|$, $\forall \lambda \in \mathbb{R}$ and $\forall v \in V$, (absolute homogeneity)

(iii) $\|u + v\| \leq \|u\| + \|v\|$, $\forall u, v \in V$. (triangle inequality)

Definition 2.2 (Normed space)**.** Let V be a vector space, and let $\|\cdot\|$ be a norm on V. The pair $(V, \|\cdot\|)$ is called a normed (vector) space.

Some authors give a slightly different, but equivalent definition of the norm.

It should be clear from the definition that the vector space structure on V is a necessity, since the axioms of being a norm rely on the existence of $0 \in V$ (the additive identity), scalar multiplication, and addition. We shall focus on the case where the underlying field is the real numbers, $\mathbb{R}$, however, an analogous theory can be developed in the case of other fields, for example, the complex numbers $\mathbb{C}$.

The idea behind defining a norm $\|\cdot\|$ is that $\|x\|$ should give us a measure of the *magnitude* of the vector x. And magnitudes (or lengths) should be non-negative. So, the norms should be non-negative valued. The next result confirms this property. Some authors directly require $\|\cdot\| : V \to [0, \infty)$, which simply gives the following proposition the status of an axiom.

Proposition 2.3 (Non-negative valued)**.** *Let $\|\cdot\| : V \to \mathbb{R}$ be a norm on V. Then, $\forall x \in V, \|x\| \geq 0$.*

Proof. Observe the following

$$0 \overset{\text{(i)}}{=} \|0\| = \|x + (-x)\| \overset{\text{(iii)}}{\leq} \|x\| + \|-x\| \overset{\text{(ii)}}{=} \|x\| + |-1|\|x\| = \|x\| + \|x\| = 2\|x\|.$$

So, $2\|x\| \geq 0$ for all $x \in V$, from which we conclude that $\|x\| \geq 0, \forall x \in V$. $\square$

Note that this proof works even when V is a vector space over a general field K.

Proposition 2.4 (Normed subspace)**.** *Let $(X, \|\cdot\|)$ be a normed space, and let Y be a vector subspace of X. Then, $(Y, \|\cdot\|)$ is also a normed space, using the same norm as X. $(Y, \|\cdot\|)$ is called the normed subspace of $(X, \|\cdot\|)$. More precisely, this proposition claims that the function $N : Y \to \mathbb{R}$ with $N(y) := \|y\|$, for all $y \in Y$, is a norm.*

Proof. Exercise. □

Exercise 2.1. Let $\|\cdot\|_a$ and $\|\cdot\|_b$ be any two norms on V. Do the following satisfy the definition of a norm

  (i) $\|\cdot\|_a + \|\cdot\|_b$

  (ii) $\lambda\|\cdot\|_a + \mu\|\cdot\|_b$, for $\lambda, \mu \in \mathbb{R}$.

  (iii) $\|\cdot\|_a \times \|\cdot\|_b$

  (iv) $\sqrt{\|\cdot\|_a}$

  (v) $\|(\|\cdot\|_b)\|_a$, with the appropriate changes in domain.

Let V be a finite dimensional vector space over $\mathbb{R}$. Then, a result from introductory linear algebra shows that $V \cong \mathbb{R}^n$. Define the family of functions $\|\cdot\|_p : \mathbb{R}^n \to \mathbb{R}$,

$$\|x\|_p := \left(\sum_{j=1}^n |x_j|^p\right)^{1/p} \quad \text{and} \quad \|x\|_\infty := \max\{x_1, \ldots, x_n\},$$

for $p \in [1, \infty)$ and $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$. Despite the suggestive notation, we have not yet claimed that these functions are valid norms.

Exercise 2.2. Show that the functions $\|\cdot\|_1$ and $\|\cdot\|_\infty$, are norms on $\mathbb{R}^n$.

Exercise 2.3. Show that the function $\|\cdot\|_p$, for $p \in (1, \infty)$, satisfies property (i) and (ii) of being a norm.

Remark. Indeed it will be shown later that $\|\cdot\|_p$ is a norm on $\mathbb{R}^n$, but it will be extremely difficult to show directly that it also satisfy the triangle inequality. In order to establish the triangle inequality we will have to develop an alternative method.

Definition 2.5. Let $(V, \|\cdot\|)$ be a normed space. We define

  (i) the open unit ball: $B_V := \{v \in V : \|v\| < 1\}$

  (ii) the closed unit ball: $\overline{B}_V := \{v \in V : \|v\| \le 1\}$

These sets depend on the norm that is being used on V.

The closed unit balls produced by the ℓ_1, ℓ_2, and ℓ_∞ norms on $\mathbb{R}^2$ are shown below.
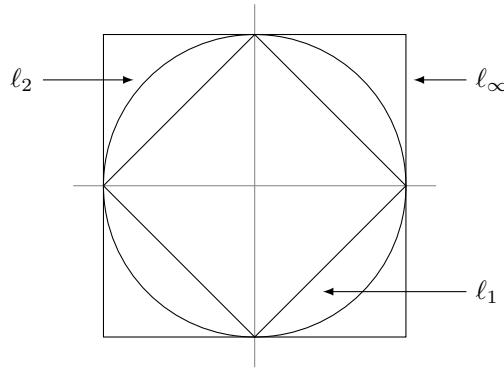
Definition 2.6 (Convex sets)**.** Let V be a vector space. A set $X \subseteq V$ is called convex if $\forall x, y \in X$, and $\forall \lambda \in (0, 1)$
$$(1 - \lambda)x + \lambda y \in X.$$

Geometrically, this is saying that X is convex if the line segment joining any every pair of points in X, lies entirely within X.

Definition 2.7 (Convex functions)**.** A function $f : A \subseteq \mathbb{R} \to \mathbb{R}$ is called convex if $\forall x, y \in A$, and $\forall \lambda \in (0, 1)$

$$f\left((1 - \lambda)x + \lambda y\right) \leq (1 - \lambda)f(x) + \lambda f(y).$$

Therefore, a function is convex exactly when its graph lies beneath (or along) the line segment joining any two points on the graph. Common examples of convex functions are x^2 and e^x on $\mathbb{R}$.

Proposition 2.8. *Let $(V, \|\cdot\|)$ be a normed space. The sets B_V and $\overline{B}_V$ are convex.*

Proof. We shall prove this for $\overline{B}_V$. Take $a, b \in \overline{B}_V$, and $\lambda \in (0, 1)$. Then,

$$\|(1 - \lambda)a + \lambda b\| \leq (1 - \lambda)\|a\| + \lambda\|b\| \leq (1 - \lambda) + \lambda = 1.$$

Therefore, $(1 - \lambda)a + \lambda b \in \overline{B}_V$. So, $\overline{B}_V$ is convex.

The proof for the convexity of B_V is analogous. $\qquad\square$

Proposition 2.9. *Suppose $f \in C^2((a, b))$ and $f \in C^1([a, b])$. If $f''(x) \geq 0$, for all $x \in [a, b]$, then f is convex on $[a, b]$.*

Proof. The proof is left as an exercise to the reader. $\qquad\square$

Lemma 2.10. *The function $f : \mathbb{R} \to [0, \infty)$ with $x \mapsto |x|$ is convex.*

Proof. To be completed. Uses the MVT. $\qquad\square$

Lemma 2.11. *The function $f : [0, \infty) \to \mathbb{R}$ with $x \mapsto x^p$ is convex, $\forall p \in [1, \infty)$.*

Proof. The given function is a polynomial, so it is infinitely differentiable, with

$$f''(x) = p(p - 1)x^{p-2}.$$

Now,

$$p \in [1, \infty), x \in [0, \infty) \implies p, (p - 1), x, x^{p-2} \geq 0.$$

Therefore, $f''(x) \geq 0$, for all x in the given domain.

Proposition (2.9) then states that $f(x) = x^p$ is convex. $\qquad\square$

Lemma 2.12. *If $f : A \to B$ and $g : B \to C$ are convex functions, then $g \circ f : A \to C$ is also convex.*

Proof. Take $x, y \in A$ and $\lambda \in (0, 1)$. Then,

$$\begin{aligned}
(1 - \lambda)(g \circ f)(a) + \lambda(g \circ f)(b) &= (1 - \lambda)g(f(a)) + \lambda g(f(b)) \\
&\geq g\left((1 - \lambda)f(a) + \lambda f(b)\right) \\
&\geq g\left(f\left((1 - \lambda)a + \lambda b\right)\right) \\
&= (g \circ f)\left((1 - \lambda)a + \lambda b\right).
\end{aligned}$$

Therefore, $(g \circ f)$ is convex. $\qquad\square$

Combining these lemmas, we obtain that $\forall p \in [1, \infty)$, $x \mapsto |x|^p$ is a convex function. In other words, $\forall x, y \in \mathbb{R}$, and $\forall \lambda \in (0, 1)$,

$$|(1 - \lambda)x + \lambda y|^p \leq (1 - \lambda)|x|^p + \lambda|y|^p.$$

Now, we consider a norm-like function that is required to satisfy all the properties except the triangle inequality. If we require the *closed balls* produced by this function to be convex, then we find that it also ends up satisfying the triangle inequality. So, the next result allows us to establish the triangle inequality by checking the convexity of the induced closed balls.

Theorem 2.13. *Let V be a vector space, and consider a function $N : V \to [0, \infty)$ with*

(i) $N(v) = 0 \iff v = 0$, for $v \in V$,

(ii) $N(\lambda v) = |\lambda| N(v)$, $\forall \lambda \in \mathbb{R}$ and $\forall v \in V$,

(iii) $B := \{v \in V : N(v) \leq 1\}$ is convex.

Then,

$$N(x + y) \leq N(x) + N(y)$$

i.e. N satisfies the triangle inequality, and therefore, f is a norm on V.

Proof. Firstly if $N(x) = 0$ then $x = 0$, and

$$N(x + y) = N(y) = N(x) + N(y).$$

Similarly, for $N(y) = 0$, we have $y = 0$, so

$$N(x + y) = N(x) = N(x) + N(y).$$

Hence, assume $N(x), N(y) \geq 0$, and pick $x, y \in V$. Then, $x/N(x), y/N(y) \in V$ by closure under scalar multiplication. Moreover, $x/N(x), y/N(y) \in B$, since

$$N\left(\frac{x}{N(x)}\right) = \left|\frac{1}{N(x)}\right| N(x) = \frac{N(x)}{N(x)} = 1.$$

Similarly for $y/N(y)$.

By convexity of B we have that

$$\left(1 - \frac{N(y)}{N(x) + N(y)}\right) \frac{x}{N(x)} + \left(\frac{N(y)}{N(x) + N(y)}\right) \frac{y}{N(y)} \in B$$

$$\implies \left(\frac{N(x)}{N(x) + N(y)}\right) \frac{x}{N(x)} + \left(\frac{N(y)}{N(x) + N(y)}\right) \frac{y}{N(y)} \in B$$

$$\implies \frac{x + y}{N(x) + N(y)} \in B$$

$$\implies N\left(\frac{x + y}{N(x) + N(y)}\right) \leq 1$$

$$\implies \frac{N(x + y)}{N(x) + N(y)} \leq 1.$$

Therefore, $N(x + y) \leq N(x) + N(y)$, as required. $\qquad\square$

Corollary 2.14 (Minkowski's inequality in $\mathbb{R}^n$)**.** *For all $p \in [1, \infty)$, if $x, y \in \mathbb{R}^n$*

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p.$$

Proof. Recall that

$$\|x\|_p := \left(\sum_{j=1}^n |x_j|^p\right)^{1/p},$$

and consider the closed ball $B = \{x \in \mathbb{R}^n : \|x\|_p \leq 1\}$. Then,

$$B = \{x \in \mathbb{R}^n : \|x\|_p \leq 1\} = \{x \in \mathbb{R}^n : \|x\|_p^p \leq 1\}.$$

We aim to show that B is convex, and then use theorem (2.13) to obtain the triangle equality. Note that $\|\cdot\|_p$ already satisfies the properties (i) and (ii) of theorem (2.13).

Now, take any $x, y \in B$ and $\lambda \in (0, 1)$. Then,

$$
\begin{aligned}
\|(1-\lambda)x + \lambda y\| &= \sum_{j=1}^n |(1-\lambda)a_j + \lambda b_j^p| \\
&\leq \sum_{j=1}^n (1-\lambda)|a_j|^p + \lambda|b_j|^p \\
&= (1-\lambda)\sum_{j=1}^n |a_j|^p + \lambda\sum_{j=1}^n |b_j|^p \\
&= (1-\lambda)\|a\|_p^p + \lambda\|b\|_p^p \\
&\leq (1-\lambda) + \lambda \\
&= 1.
\end{aligned}
$$

Therefore, $\|(1-\lambda)x + \lambda y\| \leq 1$; i.e. $(1-\lambda)x + \lambda y \in B$. This confirms that B is convex. Theorem (2.13) then states that $\|\cdot\|_p$ satisfies the triangle inequality. That is

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p. \qquad \square$$

Remark. This completes the proof that $\|\cdot\|_p$, for $p \in [1, \infty)$, as defined before, fulfils all the requirements of being a norm on $\mathbb{R}^n$.

Next, we look at equivalent norms and the associated results for their closed (open) unit balls.

Definition 2.15. Let V be a vector space. A norm $\|\cdot\|_a$ on V is said to be equivalent to another norm $\|\cdot\|_b$ on V if $\exists c_1, c_2 \in \mathbb{R}$ with $0 < c_1 \leq c_2$, such that

$$c_1\|x\|_b \leq \|x\|_a \leq c_2\|x\|_b, \forall x \in V.$$

We denote this as $\|\cdot\|_a \cong \|\cdot\|_b$.

It is important that the constants c_1 and c_2 are independent of the element $x \in V$. Clearly, if $\|\cdot\|_a$ is equivalent to $\|\cdot\|_b$, then $\|\cdot\|_b$ is equivalent to $\|\cdot\|_a$. In fact this is an equivalence relation. (Exercise & Content for the appendix)

WHAT IS THE POINT OF EQUIVALENT NORMS?

Proposition 2.16. $\|\cdot\|_a \cong \|\cdot\|_b$ *if an only if* $\exists c_1, c_2 \in \mathbb{R}$ *with* $0 < c_1 \leq c_2$, *such that* $c_1\overline{B}_b \subseteq \overline{B}_a \subseteq c_2\overline{B}_b$. *Here,* $\overline{B}_j$ *is the closed unit ball generated by the norm* $\|\cdot\|_j$.

Proof. The proof is left as an exercise to the reader. $\qquad \square$

The previous proposition states that if $\|\cdot\|_a$ is equivalent to $\|\cdot\|_b$, then the closed unit ball $\overline{B}_a$ is entirely contained within an appropriately scaled version of the closed unit ball $\overline{B}_b$, and that $\overline{B}_a$ also contains an appropriately scaled version of $\overline{B}_b$. Moreover, the converse of this statement is also true.

Theorem 2.17. *All $\|\cdot\|_p$ norms on $\mathbb{R}^n$ are equivalent.*

In fact, a more general result holds for the case of $\mathbb{R}^n$: All norms on $\mathbb{R}^n$ are equivalent. As a result, we are free to choose any of these norm to study limits and continuity in $\mathbb{R}^n$, since the results will always be equivalent.

Some other topics of interest. ℓ^p spaces, L^p spaces, other norms on $\mathbb{R}^n$, inner products, norms on $\mathbb{C}^n$.

3. Metric spaces

More general than norms is the notion of a metric. The idea is that a metric is a measure of distance between any two elements in a given set; the norm, on the other hand, is a measure of magnitude of each element individually.

COMMENT ABOUT REAL ANALYSIS.

Definition 3.1 (Metric space)**.** Let X be any set. Let $d : X \times X \to \mathbb{R}$ be a function such that

(i) $d(x, y) \geq 0, \forall x, y \in X$,   (positive)

(ii) $d(x, y) = 0 \iff x = y$,   (definite)

(iii) $d(x, y) = d(y, x), \forall x, y \in X$,   (symmetric)

(iv) $d(x, z) \leq d(x, y) + d(y, z), \forall x, y, z \in X$.   (triangle inequality)

Then, d is called a metric on X, and the pair (X, d) is called a metric space.

Unlike a norm, a metric doesn't require the underlying set to be a vector space. While we will often work with sets X that carry a natural vector space structure, the metric properties will be independent from it unless the metric is being *induced* by the norm.

Proposition 3.2 (Metric induced by a norm)**.** *let $(X, \|\cdot\|)$ be a normed vector space. Then $d : X \times X \to \mathbb{R}$ with $d(x, y) := \|x - y\|$ is a metric on X. Such a metric is called the metric induced by the norm.*

Proof. The proof is left as an exercise to the reader. $\qquad\qquad\qquad\qquad\square$

The metric $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ induced by the Euclidean norm $\|\cdot\|_2$ is known as the standard (or Euclidean) metric on $\mathbb{R}^n$. When we consider $\mathbb{R}^n$ as a metric space, without specifying the metric, the underlying metric is assumed to be the standard metric.

Example 3.3. Let X be any set, and consider the function $d_s : X \times X \to \mathbb{R}$ with

$$d_s = \begin{cases} 0, & x = y \\ 1, & x \neq y \end{cases}$$

This is a metric in X, known as the discrete metric. This is often very useful for producing counterexamples.

Example 3.4. The function $d : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ with

$$d = \begin{cases} \|\mathbf{x} - \mathbf{y}\|, & \mathbf{x} \text{ and } \mathbf{y} \text{ lie on the same line through } (0, 0) \\ \|\mathbf{x}\| + \|\mathbf{y}\|, & \text{otherwise} \end{cases}$$

defines a metric on $\mathbb{R}^2$, known as the sunflower metric. (Check.)

Example 3.5. Consider the function $d : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ with

$$d = \begin{cases} |y_1 - y_2|, & \text{if } x_1 = x_2 \\ |y_1| + |x_1 - x_2| + |y_2|, & \text{otherwise} \end{cases}$$

This is known as the jungle river metric on $\mathbb{R}^2$. (Check.)

Example 3.6. A word w, of length n is an expression of the form as $w = w_1 w_2 \cdots w_n$, where w_i are symbols (called letters) from a fixed set. The Hamming distance between any two such words $x = x_1 x_2 \cdots x_n$ and $y = y_1 y_2 \cdots y_n$ is defined to be

$$d(x, y) := \sum_{j=1}^{n} \delta(x_j, y_j), \qquad \text{with} \qquad \delta(x_j, y_j) = \begin{cases} 0, & x_j = y_j \\ 1, & x_j \neq y_j \end{cases}$$

Here we check that this function satisfies the triangle inequality. Let x, y, z be any three words of length n. Then,

$$d(x, y) + d(y, z) = \sum_{j=1}^{n} \delta(x_j, y_j) + \sum_{j=1}^{n} \delta(y_j, z_j) = \sum_{j=1}^{n} \left(\delta(x_j, y_j) + \delta(y_j, z_j) \right) \equiv \sum_{j=1}^{n} \alpha_j.$$

Here, $\alpha_j \equiv \delta(x_j, y_j) + \delta(y_j, z_j)$. Note that we have $\alpha_j \in \{0, 1, 2\}$ only. Now, we consider these three cases separately (for fixed j).

$$\begin{aligned} \alpha_j = 0: \quad & x_j = y_j = z_j \implies x_j = z_j \implies \delta(x_j, z_j) = 0 \leq \alpha_j \\ \alpha_j = 1: \quad & \delta(x_j, z_j) \in \{0, 1\} \implies \delta(x_j, z_j) \leq 1 = \alpha_j \\ \alpha_j = 2: \quad & \delta(x_j, z_j) \in \{0, 1\} \implies \delta(x_j, z_j) \leq 2 = \alpha_j \end{aligned}$$

Overall, $\delta(x_j, z_j) \leq \alpha_j$ for all j. Hence,

$$d(x, z) = \sum_{j=1}^{n} \delta(x_j, z_j) \leq \sum_{j=1}^{n} \alpha_j = d(x, y) + d(y, z).$$

Example 3.7. It is easy to check that the following are all metrics on $\mathbb{Z}$,

(i) $d(n, m) = |m - n|$,

(ii) $d(n, m) = \left| m^3 - n^3 \right|$,

(iii) $d(n, m) = \left| \tan^{-1} m - \tan^{-1} n \right|$,

(iv) $d(n, m) = \frac{|m-n|}{1+|m-n|}$,

with $m, n \in \mathbb{Z}$.

There are a few different notions of equivalence for metrics. One is a generalization of the equivalence of norms, this is known as strong equivalence.

Definition 3.8 (Strong equivalence of metrics). Two metrics d and d' on X are said to be (strongly or Lipschitz) equivalent if and only if there exist $\alpha_1, \alpha_2 \in \mathbb{R}$ with $\alpha_1, \alpha_2 > 0$, such that for all $x, y \in X$

$$\alpha_1 d(x, y) \leq d'(x, y) \leq \alpha_2 d(x, y).$$

However, a more general equivalence, called the weak or topological equivalence, relies upon the topology induced by the metrics. We shall explore this further in later sections. For now, equivalence shall mean strong equivalence.

Lemma 3.9. *Let d and d' be metrics induced by $\|\cdot\|$ and $\|\cdot\|'$, respectively. Then, d and d' are equivalent if and only if $\|\cdot\|$ and $\|\cdot\|'$ are equivalent.*

We have already seen that many different metrics can be defined on a given set X. However, given a metric space (X, d), a subset $Y \subseteq X$ can be assigned a special metric called the *subspace* metric inherited from the original metric space.

This is primarily how the metric spaces $\mathbb{Z}$ and $\mathbb{Q}$ are defined; as subsets of the metric space $\mathbb{R}$ with the standard metric.

Theorem 3.10 (Subspace metric). *Let (X, d) be a metric space and $Y \subseteq X$. Then (Y, d_Y) is a metric space, with $d_Y : Y \times Y \to \mathbb{R}$ given by*

$$d_Y(x) = d(x), \qquad \forall x \in Y.$$

(Y, d_Y) is called a metric subspace of (X, d).

The definition simply states that Y uses the same metric function that X uses.

Remark. A statement like 'the metric space $[-1, 3]$' without the identification of the metric means that the subset $[-1, 3]$ is being considered as the subspace of $\mathbb{R}$.

4. Limits and convergence in metric spaces

Throughout this section, (X, d) shall represent a general metric space.

Definition 4.1. A sequence (a_n) in X, is a function $a : \mathbb{N} \to X$ with $n \mapsto a_n \in X$.

Definition 4.2. A sequence (a_n) in X, is said to converge to $\ell \in X$ as $n \to \infty$ if

$$\forall \varepsilon > 0, \exists N \in \mathbb{N} \text{ such that } n \geq N \implies d(a_n, \ell) < \varepsilon.$$

Such a sequence is called convergent in (X, d), and we write $\lim_{n \to \infty} a_n = \ell$, or simply $(a_n) \to \ell$, suppressing the underlying metric space.

This is a generalisation of the definition of the limit of a sequence in $\mathbb{R}^n$. There $\|a_n - l\|$ is used in place of $d(a_n, \ell)$. But this simply corresponds to using the Euclidean metric for the calculation of distance.

However, this generalisation suggests that the convergence of (a_n) could depend on the choice of the metric.

Example 4.3. Consider the sequence $a_n = 1/n$ defined on $\mathbb{R}$. It is easy to show that this converges to 0 under the standard metric of $\mathbb{R}$.

However, this sequence does not converge in the discrete metric, d_s. For example, pick $\varepsilon = 1/2$, then

$$d_s\left(\frac{1}{n}, \ell\right) < \varepsilon \iff d_s\left(\frac{1}{n}, \ell\right) = 0 \iff 1/n = \ell, \forall n > N$$

for some $N \in \mathbb{N}$. This is not possible. Therefore, $1/n$ is not convergent in $(\mathbb{R}, d_s)$.

Exercise 4.1. Suppose (a_n) is a convergent sequence in a metric space X, using the discrete metric. What can be concluded about (a_n)?

Lemma 4.4. *Every convergent sequence (a_n) in (X, d) has a unique limit.*

Proof. Suppose (a_n) converges to two limits $\ell_1 \neq \ell_2$. Let $\varepsilon = \frac{1}{2} d(\ell_1, \ell_2)$. Then, $\varepsilon > 0$. So, there exist $N_1, N_2 \in \mathbb{N}$ such that

$$d(a_n, \ell_1) < \frac{\varepsilon}{2} \qquad \forall n \geq N_1$$
$$d(a_n, \ell_2) < \frac{\varepsilon}{2} \qquad \forall n \geq N_2$$

Let $N = \max\{N_1, N_2\}$. Then, for all $n \geq N$,

$$2\varepsilon = d(\ell_1, \ell_2) \leq d(\ell_1, a_n) + d(a_n, \ell_2) < \varepsilon.$$

This is a contradiction. So, $\ell_1 = \ell_2$. $\qquad\square$

Proposition 4.5. *Let d and d' be two equivalent metrics on X. A sequence (a_n) converges in X with respect to d if and only if it converges with respect to d'.*

This result clarifies an important point: equivalent norms and metrics have the same convergence properties.

Now, we consider a special class of sequences called Cauchy sequences.

Definition 4.6 (Cauchy sequence)**.** A sequence in X is called a Cauchy sequence if

$$\forall \varepsilon > 0, \exists N \in \mathbb{N} \text{ such that } n, m \geq N \implies d(a_n, a_m) < \varepsilon.$$

Theorem 4.7. *In a metric space, every Cauchy sequence is bounded.*

Proof. Let (a_n) be a Cauchy sequence. Then, there exists $N \in \mathbb{N}$ such that

$$d(x_m, x_n) < 1 \qquad \forall m, n \geq N.$$

In particular, $d(x_m, x_N) < 1$ for all $m \geq N$.

There are only finitely many terms preceeding a_N. Set

$$R = \max_{1 \leq j \leq N} d(x_j, x_N).$$

So, for all $m \leq N$, $d(x_j, x_N) \leq R$. Therefore, $d(x_j, x_N) \leq R + 1$, for all $j \in \mathbb{N}$. $\qquad \square$

Theorem 4.8. *Let (a_n) be a Cauchy sequence with a convergent subsequence $a_{n_k} \to a$. Then, $a_n \to a$ also.*

Proof. Let (a_n) be a Cauchy sequence, and suppose (a_{n_k}) is a convergent subsequence with the limit a. Then, for all $\varepsilon > 0$, there exist $N_1, N_2 \in \mathbb{N}$ such that

$$d(a_m, a_n) < \varepsilon/2 \qquad \forall m, n \geq N_1$$
$$d(a_{n_k}, a) < \varepsilon/2 \qquad \forall n_k \geq N_2.$$

Set $N = \max\{N_1, N_2\}$ and fix a $n_k \geq N$. Then, for all $m \geq N$

$$d(a_m, x) \leq d(a_m, a_{n_k}) + d(a_{n_k}, x) < \varepsilon.$$

This means that $a_m \to a$ as $m \to \infty$. $\qquad \square$

The similarity between the definition of convergent and Cauchy sequences motivates the following important result

Theorem 4.9. *Every convergent sequence in (X, d) is a Cauchy sequence.*

Proof. Let (a_n) be a convergent sequence in X with the limit ℓ. Then, for $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that
$$d(a_n, \ell) < \frac{\varepsilon}{2} \qquad \forall n \geq N.$$

So, for all $n, m \geq N$ we have

$$d(a_n, a_m) \leq d(a_n, \ell) + d(\ell, a_m) < \varepsilon.$$

Therefore, (a_n) is a Cauchy sequence. $\qquad \square$

It is crucial to note that the converse of this statement is not true in general.

Definition 4.10. A metric space (X, d) is called complete if every Cauchy sequence (a_n) in X converges to a limit ℓ in X.

Theorem 4.11 (Bolzano–Weierstrass)**.** *Every bounded sequence in $\mathbb{R}^n$ has a convergent subsequence.*

Proposition 4.12. $\mathbb{R}$ *with the standard metric is complete.*

Proof. Let (a_n) be a Cauchy sequence in $\mathbb{R}$. Then, (a_n) is bounded. This implies that it has a convergent subsequence, with the limit $a \in \mathbb{R}$. Therefore, (a_n) itself converges to $a \in \mathbb{R}$. This shows that $\mathbb{R}$ is (Cauchy) complete. $\qquad\square$

Proposition 4.13. $\mathbb{R}^d$ *with the standard metric is complete.*

Proposition 4.14. *The $\mathbb{Q}$ with the standard metric inherited from $\mathbb{R}$ is not complete.*

Proof. We need to find a sequence of rational numbers, that is also Cauchy, but whose limit is not in $\mathbb{Q}$. This is left as an exercise. $\qquad\square$

This result highlights the necessity of the limit being *inside* the metric space. If the sequence 'converges' to a value that lies *outside* the metric space, then (by definition) the given sequence is not convergent.

Exercise 4.2. Show that the open interval $(0,1) \subset \mathbb{R}$ is not complete under the usual metric inherited from $\mathbb{R}$. [Hint: find a sequence in $(0,1)$ that is Cauchy but not convergent in $(0,1)$.]

Some other topics of interest. Banach space, Hilbert space, completions.

5. Open sets & closed sets in metric spaces

Throughout this section, (X, d) shall represent a general metric space.

Definition 5.1 (Open and closed balls)**.** An open ball of radius $r > 0$, centred at $a \in X$ is a set $B_r(a) = \{x \in X \,|\, d(x,a) < r\}$. A closed ball, likewise, is a set $\overline{B}_r(a) = \{x \in X \,|\, d(x,a) \leq r\}$.

This leads us to define the notion of an *open (closed) set*; one of the most important concepts in this course.

Definition 5.2 (Open and closed sets)**.** A set $A \subseteq X$ is called open if $\forall x \in A$, $\exists \varepsilon > 0$, such that $B_\varepsilon(x) \subseteq A$. A set F is called closed if and only if its complement in X is open (i.e. iff $X - F$ is open).

It is very important to note that if a set is *not* open, then it is not necessarily closed. In particular, a set can be either open, closed, both open and closed, or neither open nor closed.

MORE EXAMPLES

Example 5.3. The single point set $\{x\}$, with $x \in \mathbb{R}$ is not open and closed in the metric space $\mathbb{R}$, using the standard metric. However, in the discrete metric $\{x\}$ is open.

Lemma 5.4. *The open ball is open. The closed ball is closed.*

Theorem 5.5. *The sets $\emptyset$ and X are always both open and closed in any metric space.*

Proof. By definition, $\emptyset$ is open if

$$\forall x \in X \,(x \in \emptyset \implies \exists \varepsilon > 0(x \in B_\varepsilon(x) \subseteq \emptyset)).$$

But $x \in \emptyset$ is always false, so this statement is vacuously true. Thus, $\emptyset$ is open.

Next, X is open if

$$\forall x \in X \,(x \in X \implies \exists \varepsilon > 0(x \in B_\varepsilon(x) \subseteq X)).$$

But both $x \in X$ and $B_\varepsilon(x) \subseteq X$ are always true. So, this statement is always true. Thus, X is open.

Now, X is open, so $\emptyset = X - X$ is closed. Similarly, $\emptyset$ is open, so $X = X - \emptyset$ is closed. $\quad\square$

The arbitrary union, and the finite intersection of open sets is open.

Theorem 5.6. *The following statements are true.*

  (i) Let A be any indexing set. If U_α is open $\forall \alpha \in A$, then $\bigcup_{\alpha \in A} U_\alpha$ is open.

  (ii) If U_j is open for all $1 \leq j \leq N$, then $\bigcap_{j=1}^N U_j$ is open.

Proof. TBC. $\qquad\square$

However, the situation is reversed for closed sets.

Theorem 5.7. *The following statements are true.*

  (i) Let A be any indexing set. If F_α is closed $\forall \alpha \in A$, then $\bigcap_{\alpha \in A} F_\alpha$ is closed.

  (ii) If F_j is closed for all $1 \leq j \leq N$, then $\bigcup_{j=1}^N F_j$ is closed.

Proof. We apply the previous theorem to obtain these results.

  (i) Let F_α be closed, then $X - F_\alpha$ is open for all $\alpha \in A$. Then,

$$X - \bigcap_{\alpha \in A} F_\alpha = \bigcup_{\alpha \in A} (X - F_\alpha)$$

  is an arbitrary union of open sets. So, it is open. Therefore, $\bigcap_{\alpha \in A} F_\alpha$ is closed.

  (ii) Let F_j be closed, then $X - F_j$ is open for all $1 \leq j \leq n$. Then,

$$X - \bigcup_{j=1}^N F_j = \bigcap_{j=1}^N (X - F_j)$$

  is a finite intersection of open sets. So, it is open. Therefore, $\bigcup_{j=1}^N F_j$ is closed.

$\qquad\square$

Example 5.8. Consider the infinite collection $\{(-1/n, \frac{1}{n}) | n \in \mathbb{N}_+\}$ of open intervals in $\mathbb{R}$. The intersection of all of these sets is $\bigcap_{n \in \mathbb{N}_+} (-1/n, 1/n) = \{0\}$, and this is not open (and closed) in $\mathbb{R}$.

This shows that the intersection of infinitely many open sets is not guaranteed to be open.

Exercise 5.1. Find an infinite collection of closed sets in $\mathbb{R}$ whose union *is not* closed in $\mathbb{R}$. Find another such collection whose union *is* closed in $\mathbb{R}$.

The terminology of open sets allows us to express the definition of convergence in the terminology of open sets. This will carry over into topological spaces.

Lemma 5.9. *For a sequence $x_n \to x$ if and only if for all open sets U containing x, there exists $N \in \mathbb{N}$ such that $x_n \in U$, for all $n \geq N$.*

Proof. Suppose $x_n \to x$. Take any U open in X, with $x \in U$. Then, $\exists \varepsilon > 0$ such that $B_\varepsilon(x) \subseteq U$.

Now, $x_n \to x$ means that $\exists N \in \mathbb{N}$ such that $d(x_n, x) < \varepsilon$, for all $n \geq N$. Therefore, $x_n \in B_\varepsilon(x) \subseteq U$, for all $n \geq N$. So, for all $n \geq N$, $x_n \in U$.

Conversely, for all $\varepsilon > 0$ consider $B_\varepsilon(x)$. This is open and contains x. So, $\exists N \in \mathbb{N}$ such that $x_n \in B_\varepsilon(x)$ for all $n \geq N$. Therefore, $d(x_n, x) < \varepsilon$ for all $n \geq N$.

Overall, $\forall \varepsilon > 0$, $\exists N \in \mathbb{N}$ such that $d(x_n, x) < \varepsilon$, for all $n \geq N$; i.e. $x_n \to x$. $\qquad\square$

Proposition 5.10. *Take subset $F \subseteq X$ and a convergent sequence $x_n \to x \in X$ such that $(x_n) \in F$ for all n. Then, F is closed if and only if $x \in F$.*

Proof. Suppose F is closed, and $x_n \in F$ for all n, and $x_n \to x$. Assume that $x \notin F$. Then, $x \in X - F$, and $X - F$ is open. So, $\exists N \in \mathbb{N}$ such that $x_n \in X - F$ for all $n \geq N$. This contradicts the criteria that $x_n \in F$ for all n. So, $x \in F$ also.

Conversely, suppose F is not closed. Then, $X - F$ is not open; i.e. there exists some $x \in X - F$ such that there is no $\varepsilon > 0$ with $B_\varepsilon(x) \subseteq X - F$.

Then, for each $k \in \mathbb{N}$, there exists some $x_k \in B_\varepsilon(x)$ such that $x_k \notin X - F$; i.e. $x_k \in F$. Then, $x_k \to x$. But then $x \in F$ by the hypothesis of the proposition, contradicting $x \in X - F$. So, F must be closed. $\qquad\square$

The following results establish a neat connection between the completeness and closedness of a subset of X.

Proposition 5.11. *Consider $S \subset X$ with the subspace metric. If S is complete then S is a closed subset of X.*

Proof. Suppose that $(x_n) \in S$ with $x_n \to x$. Then, (x_n) is a Cauchy sequence in S. So, it converges to some $y \in S$. Since, $x_n, y \in S$, we obtain that $d_S(x_n, y) = d(x_n, y)$. So, $x_n \to y$ in X. Then, by the uniqueness of the limit $y = x$. So, $x \in S$. Therefore, S is closed. $\qquad\square$

Example 5.12. We can immediately apply this result to conclude that $\mathbb{Z} \subseteq \mathbb{R}$ is closed.

Proposition 5.13. *Consider $S \subset X$ with the subspace metric. If X is complete and S is closed then S is complete.*

Proof. If (x_n) is Cauchy in S then (x_n) is also Cauchy in X, so it converges to some $x \in X$. Now, S is closed, so $x \in S$. Therefore, $x_n \to x$ in S; i.e. S is complete. $\qquad\square$

The concept of open (closed) sets is important because many of the characteristics studied in analysis can be expressed in terms of open sets. For example, the convergence of a sequence, which intuitively relies on the metric d, was neatly expressed in terms of open sets. This encourages us to define a weaker notion of equivalence for metrics.

Definition 5.14. Two metrics d and d' are called topologically equivalent, if the set of open sets of d coincides with the set of of open sets of d'.

From now only equivalence of metrics shall mean 'topological equivalence.' Indeed, it is easy to show that if two metrics are strongly equivalent, then they are also topologically equivalent.

Example 5.15. The metrics $d(x, y)$ and $d'(x, y) \coloneqq \max(1, d(x, y))$ are topologically equivalent, but not strongly equivalent.

It follows immediately from this definition that if d and d' are topologically equivalent, and a sequence (x_n) converges in (X, d), then it also converges in (X, d'). The proof is left as an exercise.

6. Continuity in metric spaces

Throughout this section, (X, d) shall represent a general metric space.

Definition 6.1. Let (X, d_X) and (Y, d_Y) be metric spaces. Then $f : X \to Y$ is said to be continuous at $a \in X$ if

$$\forall \varepsilon > 0, \exists \delta > 0, \text{ such that } d_X(x, a) < \delta \implies d_Y(f(x), f(a)) < \varepsilon.$$

Also, we say that f is continuous on X if it is continuous for all $x \in X$.

Example 6.2. For every $a \in X$, the function $f : X \to \mathbb{R}$ with $f(x) = d_X(x, a)$ is continuous.

Firstly, note that the standard metric on $\mathbb{R}$ is simply the absolute value function $|\cdot|$. Therefore,

$$d(f(x), f(a)) < \varepsilon \quad \Longleftrightarrow \quad |d_X(x, a) - d_X(a, a)| < \varepsilon$$
$$\Longleftrightarrow \quad d_X(x, a) < \varepsilon.$$

So, given any $\varepsilon > 0$, let $\delta = \varepsilon > 0$. Then, $d_X(x, a) < \delta$ implies $d(f(x), f(a)) < \varepsilon$.

Continuity of a function can also be expressed in terms of the behaviour of a convergent sequence under the given function. While this result will not be actively used later, it helps us to think about continuity as a property about bounded variation of the images of the neighbouring points in the domain. This point of view becomes tremendously useful when we begin studying topology; continuous function can have apparent jumps in their graphs.

Lemma 6.3. *If $a_n \to a$ in X and $f : (X, d_X) \to (Y, d_Y)$ is continuous at a then $f(a_n) \to f(a) \in Y$.*

Next, we verify that the sum, product and quotient of two continuous functions is also continuous, when it is well-defined. This of course requires the co-domain to have a notion of addition, multiplication or division, respectively.

Lemma 6.4. *Let $f, g : X \to \mathbb{R}$ be continuous. Then, the following functions are all continuous*

(i) $(f + g)(x) \coloneqq f(x) + g(x)$,

(ii) $(fg)(x) \coloneqq f(x)g(x)$,

(iii) $(f/g)(x) \coloneqq f(x)/g(x)$, *provided $g(x) \neq 0$.*

(Note that we fixed the co-domain to be $\mathbb{R}$ in order to make sense of addition and multiplication.)

The same result holds of the composition of two continuous functions also. We shall state and prove it later using an equivalent formulation of continuity.

Before proceeding further, we explore a few examples of continuous and discontinuous functions. LONG EXAMPLE: IMAGE OF OPEN IS NOT ALWAYS OPEN.

Theorem 6.5. *A function $f : (X, d_X) \to (Y, d_Y)$ is continuous if and only if for every open set $V \subseteq Y$, $f^{-1}(V)$ is open in X.*

Proof. Suppose that f is continuous. Take any open set $V \subset Y$. If the preimage of V, $f^{-1}(V)$, is empty then it is open, and there is nothing to show. So, assume that $f^{-1}(V)$ is not empty. Take any $x \in f^{-1}(V) \subseteq X$. Then $f(x) \in V$.

V is open, so there exists $\varepsilon > 0$ such that $B_{Y,\varepsilon}(f(x)) \subseteq V$. Now,

$$y \in B_{Y,\varepsilon}(f(x)) \implies y \in V \iff d_Y(f(x), y) < \varepsilon \implies y \in V.$$

As f is continuous, for this $\varepsilon > 0$, there exists $\delta > 0$ such that $d_X(x, a) < \delta$ implies that $d_Y(f(x), f(a)) < \varepsilon$. In other words,

$$a \in B_{X,\delta}(x) \implies f(a) \in B_{Y,\varepsilon}(f(x)) \implies f(a) \in V.$$

Therefore, $B_{X,\delta}(x) \subseteq f^{-1}(V)$. So, $f^{-1}(V)$ is open.

For the converse, take $x \in X$ and some $\varepsilon > 0$. Then, $B_{Y,\varepsilon}(f(x))$ is open in Y, so $f^{-1}(B_{Y,\varepsilon}(f(x)))$ is open in X.

Since, $x \in B_{X,\delta}(x)$, we have $B_{X,\delta}(x) \subseteq f^{-1}(B_{Y,\varepsilon}(f(x)))$ for some $\delta > 0$. In other words,

$$d_X(x, a) < \delta \implies d_Y(f(x), f(a)) < \varepsilon. \qquad \square$$

This has an immediate, and curious corollary.

Corollary 6.6. *Let (X, d_s) be the metric space with the discrete metric, and (Y, d) be any metric space. Then every function $f : X \to Y$ is continuous.*

Proof. Let $f : X \to Y$ be any function. We note that since X has the discrete metric, every subset of X is open in X.

Let $U \subseteq Y$ be any open subset of Y. Now, either $f^{-1}(U)$ is empty. Or it is a non-empty open subset $A \subseteq X$. If $f^{-1}(U) = \emptyset$, then it is open in X. Moreover, if $f^{-1}(U) = A \subseteq X$, then it is also open in X, because every subset of X is open. Therefore, $f^{-1}(U)$ is always open, showing that f is continuous. $\qquad\square$

In the previous theorem, we chose to express continuity in terms of open sets. However, we could have chosen to use closed sets to characterise this also.

Lemma 6.7. *If $f : X \to Y$ and $A \subseteq Y$, then $f^{-1}(Y - A) = X - f^{-1}(A)$.*

Corollary 6.8. *A function $f : (X, d_X) \to (Y, d_Y)$ is continuous if and only if for every closed set $V \subseteq Y$, $f^{-1}(V)$ is closed in X.*

Proof. If $f : X \to Y$ is continuous and $F \subseteq Y$ is closed then

$$X - f^{-1}(F) = f^{-1}(Y - F).$$

And this is open because $Y - F$ is open.

Conversely, if for every closed set $F \subseteq Y$, $f^{-1}(F)$ is closed, then taking any open subset $V \subseteq Y$ we get

$$f^{-1}(Y - V) = X - f^{-1}(V),$$

is closed in X. So, $f^{-1}(V)$ is open in X. $\qquad\square$

Now, we can easily show the continuity of the composition of two continuous functions.

Proposition 6.9. *Let X, Y, Z be metric spaces. Let $f : X \to Y$, and $g : Y \to Z$ be continuous functions. Then, $g \circ f : X \to Z$ is also continuous.*

Proof. Take $U \subseteq Z$ open. Then $g^{-1}(U) \subseteq Y$ is open. So, $f^{-1}(g^{-1}(U)) \equiv (g \circ f)^{-1}(U)$ is open in X. $\qquad\square$

PRACTICALLY IMPORTANT: USING CTY TO SHOW THAT SETS ARE OPEN OR CLOSED.

CONCEPTUALLY IMPORTANT: ABOUT TOPOLOGICALLY EQUIVALENT METRICS.

Before closing this section, we define a few very important classes.

An isometry is a distance-preserving map between metric spaces.

Definition 6.10 (Isometry)**.** An isometry between X and Y is a bijective function $f : X \to Y$ such that

$$d_Y(f(x), f(y)) = d_X(x, y).$$

If an isometry exists between X and Y, then they are called isometric (to each other). It is an easy exercise to show that an isometry is always continuous. (This follows from the previously studied example that $x \mapsto d(x, a)$ is continuous for all $a \in X$.)

Next, if we drop the relationship between the metrics, and instead require bi-continuity explicitly, then we arrive at the notion of a homeomorphism. This will play a central role in the study of topological spaces.

Definition 6.11 (Homeomorphism)**.** A homeomorphism between X and Y is a bijective function $f : X \to Y$ such that both f and f^{-1} are continuous.

Exercise. Show that

(i) if $f : X \to Y$ is an isometry then it is continuous.

(ii) it is possible that $f : X \to Y$ is bijective and continuous, while f^{-1} fails to be continuous. (Hint: equip X with the discrete metric.)

(iii) every isometry is a homeomorphism.

(iv) not every homeomorphism is an isometry.

(v) the constant map, and the identity map is an isometry.

7. Topological spaces

Definition 7.1. A topology on a set X is a set $T \subseteq \mathcal{P}(X)$ of subsets of X, such that

(T1) $\emptyset, X \in T$

(T2) The union of every arbitrary collection of elements of T, is an element of T.

(T3) The intersection of every finite collection of elements of T, is an element of T.

The pair (X, T) is called a topological space. The elements of T are called open sets.

EXPLANATION OF (T2) AND (T3) AND THE NAME OPEN SETS.

CLOSED IF ITS COMPLEMENT IS OPEN.

ALTERNATIVE DEFINITION OF TOPOLOGY IN TERMS OF CLOSED SETS.

TOPOLOGY INDUCED BY THE METRIC.

MORE EXAMPLES.

BASIS AND SUB-BASIS.

SUBSPACE TOPOLOGY AND PRODUCT TOPOLOGY.

NEIGHBOURHOOD, CLOSURE, INTERIOR, BOUNDARY. AND SEVERAL LEM-MAS.

LIMIT POINTS AND ISOLATED POINTS.

LARGE AND SMALL SUBSETS. DENSE, NOWHERE DENSE, MEAGRE.

EXAMPLE: CANTOR SET.

HAUSDORFF SPACES. OTHER SEPARATION AXIOMS. UNIQUE LIMITS OF SEQUENCES. METRISABILITY.

MAPS BETWEEN TOPOLOGICAL SPACES. AND CONTINUITY.

TOPOLOGICAL INVARIANCE.

COMPACTNESS.

CONNECTEDNESS.

References

[1] W. A. Sutherland, Introduction to Metric & Topological Spaces, 2nd Edition, Oxford University Press, Reading, Massachusetts, 2009.