

Exercise 5.1. Let $X = \mathbb{R}^n$ and define the function $d_{infy} : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ as

$$d_{\infty}(x, y) = \max\{|x^1 - y^1|, \dots, |x^n - y^n|\}.$$

Show that d_{∞} is a metric on \mathbb{R}^n .

Solution: We must verify the three properties M1-M3.

M1: By the properties of the modulus function, for all $x \in \mathbb{R}$, $|x| \geq 0$. This implies that $d_{\infty}(x, y) \geq 0$. Moreover, for every $x \in \mathbb{R}$, $|x| = 0$ iff $x = 0$. Therefore, $d_{\infty}(x, y) = 0$ iff $x^i = y^i$ for all $i = 1, 2, \dots, n$ iff $x = y$.

M2: Since $|-x| = |x|$, we have

$$d_{\infty}(x, y) = \max\{|x^1 - y^1|, \dots, |x^n - y^n|\} = \max\{|y^1 - x^1|, \dots, |y^n - x^n|\} = d_{\infty}(y, x).$$

M3: Let

$$x = (x^1, x^2, \dots, x^n), \quad y = (y^1, y^2, \dots, y^n), \quad z = (z^1, z^2, \dots, z^n)$$

be arbitrary elements in \mathbb{R}^n . By the triangle inequality for the modulus, for every $i = 1, 2, \dots, n$, we have

$$|x^i - z^i| \leq |x^i - y^i| + |y^i - z^i|.$$

For every $k \in \{1, 2, \dots, n\}$ we have

$$\begin{aligned} |x^k - z^k| &\leq |x^k - y^k| + |y^k - z^k| \\ &\leq \max\{|x^1 - y^1|, \dots, |x^n - y^n|\} + \max\{|y^1 - z^1|, \dots, |y^n - z^n|\} \\ &= d_{\infty}(x, y) + d_{\infty}(y, z). \end{aligned}$$

This implies that

$$d_{\infty}(x, z) = \max\{|x^1 - z^1|, \dots, |x^n - z^n|\} \leq d_{\infty}(x, y) + d_{\infty}(y, z).$$

Alternatively, the last step for the proof of property M3, can be given as follows. First note that if A and B are finite sets of real numbers, we have

$$\max(A + B) \leq \max A + \max B.$$

Therefore,

$$\begin{aligned} d_{\infty}(x, z) &= \max\{|x^1 - z^1|, \dots, |x^n - z^n|\} \\ &= \max\{|x^1 - y^1 + y^1 - z^1|, |x^2 - y^2 + y^2 - z^2|, \dots, |x^n - y^n + y^n - z^n|\} \\ &\leq \max\{|x^1 - y^1| + |y^1 - z^1|, |x^2 - y^2| + |y^2 - z^2|, \dots, |x^n - y^n| + |y^n - z^n|\} \\ &\leq \max\{|x^1 - y^1|, \dots, |x^n - y^n|\} + \max\{|y^1 - z^1|, \dots, |y^n - z^n|\} \\ &= d_{\infty}(x, y) + d_{\infty}(y, z). \end{aligned}$$

Exercise 5.2. Show that each of the following functions is a metric on \mathbb{R} :

- (i) $d(x, y) = |x^3 - y^3|$, (here x^3 means x raised to power 3)
- (ii) $d(x, y) = |e^x - e^y|$,
- (iii) $d(x, y) = |\tan^{-1}(x) - \tan^{-1}(y)|$.

Which property of the maps $x \mapsto x^3$, $x \mapsto e^x$, and $x \mapsto \tan^{-1}(x)$ makes these functions a metric.

Solution: Let $f(x)$ stand for any of the functions $x \mapsto x^3$, $x \mapsto e^x$, and $x \mapsto \tan^{-1}(x)$. By the properties of the modulus function, we immediately obtain $d(x, y) \geq 0$, and $d(x, y) = d(y, x)$. Also, by the inequalities

$$d(x, y) = |f(x) - f(y)| \leq |f(x) - f(z)| + |f(z) - f(y)| = d(x, z) + d(z, y),$$

we obtain the triangle inequality for the functions d in each case.

There remains to see that $d(x, y) = 0$ iff $x = y$. Clearly, if $x = y$, $d(x, x) = 0$. The opposite implication follows from the fact that f is injective in all the three cases. That is, if $f(x) = f(y)$ for some x and y in \mathbb{R} , we must have $x = y$.

This exercise shows that there are many metrics on \mathbb{R} , as there are many injective maps from \mathbb{R} to \mathbb{R} . Note that the continuity of f is not required here.

Exercise 5.3. Assume that $a < b$ are real numbers, and $h : (a, b) \rightarrow (0, \infty)$ is a continuous function. For x and y in (a, b) , we define

$$d_h(x, y) = \int_{\min\{x, y\}}^{\max\{x, y\}} h(t) dt.$$

Show that d_h is a metric on (a, b) .

Solution: M2: Since $\{x, y\} = \{y, x\}$ as sets, by the definition of d_h , we immediately see that $d_h(x, y) = d_h(y, x)$. Therefore, without loss of generality, below we assume that $x \leq y$.

M1: For real numbers $x \leq y$ and a function $h \geq 0$, the Riemann integral satisfies $\int_x^y h(t) dt \geq 0$. Moreover, by the definition of integral, if $x = y$ we have $\int_x^y h(t) dt = 0$. On the other hand, by a lemma proved in the typed lectures, if $h > 0$ and $x < y$, we must have $\int_x^y h(t) dt > 0$. Since $h > 0$, this implies that if $\int_x^y h(t) dt = 0$ we must have $x = y$. Therefore, $d_h(x, y) = 0$ iff $x = y$.

M3: Let x, y and z be arbitrary real numbers. Without loss of generality, assume that $x \leq y$. Recall from the properties of the Riemann integral that if $x \leq z \leq y$, we have

$$\int_x^y h(t) dt = \int_x^z h(t) dt + \int_z^y h(t) dt.$$

This implies that for all real numbers $x \leq z \leq y$, we have

$$d_h(x, y) \leq d_h(x, z) + d_h(z, y).$$

If $z \notin [x, y]$, we must either have $z \leq x$ or $z \geq y$. In the first case, we have $[x, y] \subset [z, y]$, and hence

$$\int_x^y h(t) dt \leq \int_z^y h(t) dt = \int_z^x h(t) dt + \int_x^y h(t) dt$$

and in the second case we have $[x, y] \subseteq [x, z]$, and hence

$$\int_x^y h(t)dt \leq \int_x^z h(t)dt = \int_x^y h(t)dt + \int_y^z h(t)dt.$$

Each of these inequalities imply that

$$d_h(x, y) \leq d_h(x, z) + d_h(z, y).$$

Exercise 5.4. Consider the function $g : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ defined as

$$g(x, y) = |x - y|^2.$$

Show that g is not a metric on \mathbb{R} .

Solution: It is sufficient to show that one of the properties of the metric does not hold. Consider the three points 2, 3, 4. Then,

$$g(2, 4) = 4 \not\leq 1 + 1 = g(2, 3) + g(3, 4).$$

This shows the triangle inequality does not hold for the three points 2, 3, 4.

Another counter example is given by the three points 0, 10, 20 in \mathbb{R} , as

$$g(0, 20) = 400 \not\leq 200 = 100 + 100 = g(0, 10) + g(10, 20).$$

Exercise 5.5. Let $X = \mathbb{R}^2$, and define $d_{\text{rail}} : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$ as

$$d_{\text{rail}}(x, y) = \begin{cases} \|x - y\| & \text{if } x = ky \text{ for some } k \in \mathbb{R} \\ \|x\| + \|y\| & \text{otherwise} \end{cases}$$

Show that d_{rail} is a metric on \mathbb{R}^2 .

This is called the British rail metric. The intuition behind this metric is that if two towns are on the same rail line, then we travel between them, but if the towns are on distinct lines, we travel via London (represented as the origin in \mathbb{R}^2).

Solution: The properties $d_{\text{rail}}(x, y) \geq 0$, $d_{\text{rail}}(x, y) = 0$ iff $x = y$, and $d_{\text{rail}}(x, y) = d_{\text{rail}}(y, x)$ easily follow from the properties of the norm. We need to show the triangle inequality for d_{rail} . Let x, y , and z be arbitrary points in \mathbb{R}^2 . We consider few cases below.

1) Assume that there is $k \in \mathbb{R}$ such that $x = ky$.

$$d_{\text{rail}}(x, y) = \|x - y\| \leq \|x - z\| + \|z - y\| \leq d_{\text{rail}}(x, z) + d_{\text{rail}}(z, y),$$

since we always have $\|a - b\| \leq \|a\| + \|b\|$.

2) Assume that for all $k \in \mathbb{R}$ we have $x \neq ky$. In particular, $x \neq 0$. There are several cases to look at in this case.

(i) $z = 0$. We have

$$d_{\text{rail}}(x, y) = \|x\| + \|y\| = d_{\text{rail}}(x, 0) + d_{\text{rail}}(0, y)$$

(ii) There is $m \in \mathbb{R} \setminus \{0\}$ such that $z = my$. Then, for all $s \in \mathbb{R}$, we have $z \neq sx$, otherwise $my = sx$, and (as $s \neq 0$) $x = (m/s)y$, which is a contradiction. In particular,

$$d_{\text{rail}}(x, y) = \|x\| + \|y\| \leq \|x\| + \|z - y\| + \|z\| = d_{\text{rail}}(x, z) + d_{\text{rail}}(z, y).$$

(iii) There is $l \in \mathbb{R} \setminus \{0\}$ such that $z = lx$. This is similar to case (ii), as one may switch x and y in that proof.

(iv) For all $m \in \mathbb{R}$, we have $z \neq my$ and $z \neq mx$. Then,

$$d_{\text{rail}}(x, y) = \|x\| + \|y\| \leq \|x\| + \|z\| + \|z\| + \|y\| = d_{\text{rail}}(x, z) + d_{\text{rail}}(z, y).$$

Exercise 5.6. Assume that $a < b$ are real numbers. Show that each of the following functions is a norm on $C([a, b])$:

(i)

$$\|f\|_1 = \int_a^b |f(t)| dt$$

(ii)

$$\|f\|_\infty = \max_{t \in [a, b]} |f(t)|$$

(iii)

$$\|f\|_2 = \left(\int_a^b |f(t)|^2 dt \right)^{1/2}$$

Hint: to show that $\|\cdot\|_2$ is a norm, you need to use the Cauchy-Schwarz inequality and the definition of the integral as the limit of certain sums.

Solution: (i) By the properties of the Riemann integral, $\|f\|_1 \geq 0$. By a lemma in the lecture notes, $\|f\|_1 = 0$ iff $f \equiv 0$. For every $\lambda \in \mathbb{R}$, we have

$$\|\lambda f\|_1 = \int_a^b |\lambda f(t)| dt = \int_a^b |\lambda| |f(t)| dt = |\lambda| \int_a^b |f(t)| dt = |\lambda| \|f\|_1.$$

Moreover, for all f and g in $C([a, b])$, we have

$$\|f + g\|_1 = \int_a^b |f(t) + g(t)| dt \leq \int_a^b (|f(t)| + |g(t)|) dt = \int_a^b |f(t)| dt + \int_a^b |g(t)| dt,$$

which implies that $\|f + g\|_1 \leq \|f\|_1 + \|g\|_1$.

(ii) For every f in $C([a, b])$, the maximum of f on $[a, b]$ is realised, so $\|f\|_\infty$ is well-defined, and a real number. Evidently, $\|f\|_\infty \geq 0$, and $\|f\|_\infty = 0$ iff $f \equiv 0$. Moreover, for all $\lambda \in \mathbb{R}$, we have

$$\|\lambda f\|_\infty = \max_{t \in [a, b]} |\lambda f(t)| = \max_{t \in [a, b]} (|\lambda| |f(t)|) = |\lambda| \max_{t \in [a, b]} |f(t)| = |\lambda| \|f\|_\infty.$$

Finally, for all f and g in $C([a, b])$, we have

$$\begin{aligned} \|f + g\|_\infty &= \max_{t \in [a, b]} |f(t) + g(t)| \\ &\leq \max_{t \in [a, b]} (|f(t)| + |g(t)|) \\ &\leq \max_{t \in [a, b]} |f(t)| + \max_{t \in [a, b]} |g(t)| \\ &= \|f\|_\infty + \|g\|_\infty. \end{aligned}$$

(iii) Fix arbitrary functions f and g in $C([a, b])$. We note that for all $\lambda \in \mathbb{R}$, we have

$$\int_a^b (f(t) - \lambda g(t))^2 dt \geq 0.$$

This implies that

$$\int_a^b f(t)^2 dt - 2\lambda \int_a^b f(t)g(t) dt + \lambda^2 \int_a^b g(t)^2 dt \geq 0.$$

One may think of the expression on the left hand side of the above equation as a quadratic polynomial in λ . We know that if a quadratic polynomial of the above form is non-negative, then the discriminant (“ $b^2 - 4ac$ ”) must be non-positive, that is,

$$4 \left(\int_a^b f(t)g(t) dt \right)^2 \leq 4 \int_a^b f(t)^2 dt \cdot \int_a^b g(t)^2 dt.$$

This implies that for all f and g in $C([a, b])$, we have

$$\left| \int_a^b f(t)g(t) dt \right| \leq \|f\|_2 \|g\|_2.$$

The above inequality is known as the Cauchy–Schwarz inequality. It is also possible to prove the above inequality, using the definition of the integral as limits of sums, and using the Cauchy-Schwarz inequality in \mathbb{R}^n .

Using the Cauchy-Schwarz inequality, we can see that for all f and g in $C([a, b])$, we have

$$\|f + g\|_2^2 = \int_a^b |f(t) + g(t)|^2 dt = \int_a^b f(t)^2 dt + 2 \int_a^b f(t)g(t) dt + \int_a^b g(t)^2 dt \leq (\|f\|_2 + \|g\|_2)^2,$$

which implies that $\|f + g\|_2 \leq \|f\|_2 + \|g\|_2$.

The other properties for $\|\cdot\|_2$ can be proved by arguments similar to the ones for $\|\cdot\|_1$.

Exercise 5.7. Show that if V is a vector space, and $\|\cdot\| : V \rightarrow \mathbb{R}$ is a norm function, then for any $v \in V$, we must have $d_{\|\cdot\|}(0, 2v) = 2 d_{\|\cdot\|}(0, v)$. Conclude that there is no norm function on \mathbb{R}^2 which induced the discrete metric d_{disc} on \mathbb{R}^2 .

Solution: Since for every norm function, any $v \in V$ and any $\lambda \in \mathbb{R}$, we have $\|\lambda v\| = |\lambda| \|v\|$, we must have

$$d_{\|\cdot\|}(0, 2v) = \|2v\| = 2 \|v\| = 2 d_{\|\cdot\|}(0, v).$$

For the discrete metric, we have

$$d_{\text{disc}}((0, 0), (1, 1)) = d_{\text{disc}}((0, 0), (2, 2)) = 1,$$

which does not satisfy the above relation when $v = (1, 1)$.

Exercise 5.8. Let (X, d) be a metric space.

(i) Show that for every x, y , and z in X , we have

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

(ii) Show that for all x, y, z and t in X , we have

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

(iii) Show that for all x_1, x_2, \dots, x_n in X , we have

$$d(x_1, x_n) \leq d(x_1, x_2) + d(x_2, x_3) + \dots + d(x_{n-1}, x_n).$$

Solution: (i) Using the triangle inequalities

$$d(x, z) \leq d(x, y) + d(y, z), \quad d(y, z) \leq d(x, y) + d(x, z),$$

we obtain

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

which is equivalent to the the desired inequality.

(ii) Using the triangle inequality two times, we obtain

$$d(x, y) \leq d(x, z) + d(y, z) \leq d(x, z) + d(z, t) + d(y, t),$$

and

$$d(z, t) \leq d(z, x) + d(x, t) \leq d(z, x) + d(x, y) + d(y, t).$$

By adding and subtracting appropriate terms, we obtain

$$d(x, y) - d(z, t) \leq d(x, z) + d(y, t),$$

and

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

These two inequalities imply the desired inequality in part (ii).

(iii) We prove the desired statement by induction on the number of points, n . For $n = 2$ the inequality is obvious. Assume that the inequality holds for n points. For any collection of $n + 1$ points, x_1, x_2, \dots, x_{n+1} , we have

$$\begin{aligned} d(x_1, x_{n+1}) &\leq d(x_1, x_n) + d(x_n, x_{n+1}) \\ &\leq d(x_1, x_2) + d(x_2, x_3) + \dots + d(x_{n-1}, x_n) + d(x_n, x_{n+1}). \end{aligned}$$

Exercise 5.9. Let (X, d) be a metric space.

(i) Show that if $\epsilon < \delta$, then $B_\epsilon(x) \subseteq B_\delta(x)$. By an example, show that the equality may hold even if $\epsilon < \delta$.

(ii) Show that for every $x \in X$, we have

$$\bigcap_{n \in \mathbb{N}} B_{1/n}(x) = \{x\}.$$

Solution: (i) If $y \in B_\epsilon(x)$, then $d(x, y) < \epsilon$, and hence $d(x, y) < \delta$. Therefore, $y \in B_\delta(x)$. In the discrete metric on \mathbb{R} , $B_2(0) = B_3(0) = \mathbb{R}$.

(ii) It is enough to show that $\{x\} \subseteq \bigcap_{n \in \mathbb{N}} B_{1/n}(x)$ and $\bigcap_{n \in \mathbb{N}} B_{1/n}(x) \subseteq \{x\}$. Since for all $n \geq 1$ we have $x \in B_{1/n}(x)$, we conclude that $x \in \bigcap_{n \in \mathbb{N}} B_{1/n}(x)$.

Fix an arbitrary $y \in \bigcap_{n \in \mathbb{N}} B_{1/n}(x)$. Then, for every $n \geq 1$ we have $d(x, y) < 1/n$. This implies that $d(x, y) = 0$, and by the property of the metrics, we obtain $y = x$. Therefore, $y \in \{x\}$.

Exercise 5.10. (i) Show that for all x and y in \mathbb{R}^n , we have

$$d_\infty(x, y) \leq d_2(x, y) \leq \sqrt{n} \cdot d_\infty(x, y).$$

(ii) Show that for all x and y in \mathbb{R}^n , we have

$$d_\infty(x, y) \leq d_1(x, y) \leq n \cdot d_\infty(x, y).$$

(iii) Show/conclude that for all x and y in \mathbb{R}^n , we have

$$\frac{1}{\sqrt{n}} d_2(x, y) \leq d_1(x, y) \leq \sqrt{n} d_2(x, y).$$

(iv) Conclude that the metrics d_1 , d_2 and d_∞ on \mathbb{R}^n are topologically equivalent.

Solution: (i) This is the statement in Exercise 1.2, formulated in a different form.

(ii) If $x = (x^1, x^2, \dots, x^n)$ and $y = (y^1, y^2, \dots, y^n)$, we have

$$\max_{j=1, \dots, n} |x^j - y^j| \leq \sum_{j=1}^n |x^j - y^j| \leq n \max_{j=1, \dots, n} |x^j - y^j|.$$

(iii) These immediately follow from the inequalities in part (i), (ii).

(iv) We need to show that for any set $U \subseteq \mathbb{R}^n$, U is open with respect to d_1 , if and only if U is open with respect to d_2 , if and only if U is open with respect to d_∞ . Let us assume that U is open with respect to d_1 .

Fix an arbitrary $x \in U$. Since U is open with respect to d_1 , there is $r > 0$ such that

$$B_r(x, \mathbb{R}^n, d_1) \subseteq U.$$

By the right-hand side of the inequality in part (iii), we have

$$B_{r/\sqrt{n}}(x, \mathbb{R}^n, d_2) \subseteq B_r(x, \mathbb{R}^n, d_1).$$

Therefore,

$$B_{r/\sqrt{n}}(x, \mathbb{R}^n, d_2) \subseteq U.$$

Because $x \in U$ was arbitrary, this implies that U is open with respect to d_2 .

Similarly, by the right-hand side of the inequality in part (ii), we have

$$B_{r/n}(x, \mathbb{R}^n, d_\infty) \subseteq B_r(x, \mathbb{R}^n, d_1).$$

This implies that

$$B_{r/n}(x, \mathbb{R}^n, d_\infty) \subseteq U.$$

As $x \in U$ was arbitrary, this implies that U is open with respect to d_∞ .

All the other implications can be proved in a similar fashion using the other sides of the inequalities in part (ii) and (iii).

Unseen Exercise. Let $E = \{1, 2, 3, 4, 5, 6\}$, and let $\mathcal{P}(E)$ be the set of all subsets of E . Consider the metric d_{card} on $\mathcal{P}(E)$ (see typed lecture notes). Let $e = \{1, 2, 3\} \in \mathcal{E}$. What is $B_{1/2}(e)$? What is $B_1(e)$? What is $B_{3/2}(e)$?

Solution: By definition, $B_\epsilon(e)$ is the set of all points $y \in \mathcal{P}(E)$ such that $d_{\text{card}}(e, y) < \epsilon$. By definition, $d_{\text{card}}(x, y) = \text{Card}(x \Delta y)$.

Fix an arbitrary $r \in (0, 1)$. If $y \in \mathcal{P}(E)$, and $d_{\text{card}}(e, y) < r$, we must have

$$\text{Card}((e \setminus y) \cup (y \setminus e)) = \text{Card}(e \setminus y) + \text{Card}(y \setminus e) < r.$$

This is because the sets $e \setminus y$ and $y \setminus e$ are disjoint sets. The above inequality implies that

$$\text{Card}(e \setminus y) < 1, \text{ and } \text{Card}(y \setminus e) < 1.$$

The inequality on the left hand side implies that $e \setminus y = \emptyset$ and hence $e \subseteq y$. Similarly, the inequality on the right hand side implies that $y \subseteq e$. Therefore, $y = e$. On the other hand, since $r > 0$, we have $e \in B_r(e)$. Combining these together, we obtain $B_r(e) = \{e\}$. In particular, $B_{1/2}(e) = B_1(e) = e$.

By the definition of the metric d_{card} , if $y \in B_{3/2}(e)$, y may have at most one more element than the set e or at most one element less than e . Therefore,

$$B_{3/2}(e) = \{e, \{1, 2, 3, 4\}, \{1, 2, 3, 5\}, \{1, 2, 3, 6\}, \{1, 2\}, \{1, 3\}, \{2, 3\}\}.$$