2 Metrics and norms

Definition 2.1. Let $X$ be a non-empty set. A function $\varrho : X \times X \to \mathbb{R}$ is called a metric on $X$ if it satisfies

M1. $\varrho(x, y) > 0$ if $x \neq y$, and $\varrho(x, x) = 0$,

M2. $\varrho(x, y) = \varrho(y, x)$,

M3. $\varrho(x, z) \leq \varrho(x, y) + \varrho(y, z)$,

where $x, y$ and $z$ are arbitrary elements of $X$.

Example 2.2. $\mathbb{R}$ with the usual (Euclidean) metric $\varrho(x, y) = |x - y|$.

Example 2.3 (Euclidean or $\ell^2$-distance). $\mathbb{R}^n$ with the Euclidean metric

$$\varrho(x, y) = \left( \sum_{i=1}^{n} |x_i - y_i|^2 \right)^{1/2},$$

where $x_i$ and $y_i$ are coordinates of the points $x$ and $y$ respectively.

Example 2.4 (Manhattan or $\ell^1$-distance). $\mathbb{R}^n$ with the metric

$$\varrho(x, y) = |x_1 - y_1| + |x_2 - y_2| + \cdots + |x_n - y_n|.$$

Example 2.5 ($\ell^\infty$-distance). $\mathbb{R}^n$ with the metric

$$\varrho(x, y) = \max \{|x_1 - y_1|, |x_2 - y_2|, \ldots, |x_n - y_n|\}.$$

Example 2.6. $C[a, b]$, the set of all continuous real functions on $[a, b]$, with the metric

$$\varrho(f, g) = \sup_{x \in [a, b]} |f(x) - g(x)|. \tag{1}$$

Example 2.7. $B(S)$, the set of all bounded real functions on a set $S$, with the metric

$$\varrho(f, g) = \sup_{x \in S} |f(x) - g(x)|. \tag{2}$$

Note that when $S = [a, b]$, equations (1) and (2) define the same metric. However the sets on which those metrics are defined are different in the two examples. (We return to this point later.)
Example 2.8 (discrete metric). For any set $X$, define the metric $\varrho$ by

$$\varrho(x, y) = \begin{cases} 
1 & \text{if } x \neq y; \\
0 & \text{otherwise.}
\end{cases}$$

Of course, it is necessary to check that all the above examples are indeed metrics, i.e., that they satisfy Definition 2.1. Usually it is easy to check conditions M1 and M2 (symmetry), but M3 may require more work. We’ll consider just the more important or trickier examples here, leaving the others as exercises.

Consider for instance the metric $\varrho$ on $\mathbb{R}$ defined in Example 2.2. It is clear that $\varrho$ satisfies M1 and M2. For M3, we need to show that $\varrho(x, z) \leq \varrho(x, y) + \varrho(y, z)$, for all $x, y, z \in \mathbb{R}$. But this is just the assertion that $|x - z| \leq |x - y| + |y - z|$ for all $x, y, z \in \mathbb{R}$, which is the familiar “triangle inequality” for real numbers.

Example 2.3 (continued). Conditions M1 and M2 are easy to check. (Do this!) For the triangle inequality M3 we need to check that

$$\sqrt{\sum_i (x_i - z_i)^2} \leq \sqrt{\sum_i (x_i - y_i)^2} + \sqrt{\sum_i (y_i - z_i)^2},$$

for all points $x = (x_i), y = (y_i), z = (z_i)$ in $\mathbb{R}^n$. Here, and in what follows, all summations are over $i = 1, \ldots, n$.

For brevity, write $u_i$ for $x_i - y_i$ and $v_i$ for $y_i - z_i$ and note that $x_i - z_i = u_i + v_i$. The inequality we need to prove becomes

$$\sqrt{\sum_i (u_i + v_i)^2} \leq \sqrt{\sum_i u_i^2} + \sqrt{\sum_i v_i^2}.$$

Since both sides of this inequality are positive, we can square both of them to obtain an equivalent inequality

$$\sum_i (u_i + v_i)^2 \leq \sum_i u_i^2 + \sum_i v_i^2 + 2 \sqrt{\sum_i u_i^2} \sqrt{\sum_i v_i^2},$$

which simplifies to

$$\sum_i u_i v_i \leq \sqrt{\sum_i u_i^2} \sqrt{\sum_i v_i^2}.$$

But this is the well-known Cauchy-Schwarz inequality.
Example 2.7 (continued). Suppose \( f, g, h : S \to \mathbb{R} \) are bounded real-valued functions on \( S \). Since \( f \) and \( g \) are bounded, so also is \( f - g \). So 
\[
\rho(f, g) = \sup_{x \in S} |f(x) - g(x)| = 0
\]
is the supremum of a bounded set and hence well defined.

It is clear that \( \rho(f, g) \geq 0 \). Also,
\[
\rho(f, g) = 0 \iff \sup_{x \in S} |f(x) - g(x)| = 0
\iff |f(x) - g(x)| = 0, \text{ for all } x \in S
\iff f(x) = g(x), \text{ for all } x \in S.
\]
So M1 holds.

For M2, note that \( \rho(f, g) = \sup_{x \in S} |f(x) - g(x)| = \sup_{x \in S} |g(x) - f(x)| = \rho(g, f) \).

M3 requires only a little more work. Let \( \alpha \in S \) be arbitrary. By definition, \( \rho(f, g) = \sup_{x \in S} |f(x) - g(x)| \) and so \( |f(\alpha) - g(\alpha)| \leq \rho(f, g) \). Similarly, \( |g(\alpha) - h(\alpha)| \leq \rho(g, h) \). By the triangle inequality for real numbers, \( |f(\alpha) - h(\alpha)| \leq |f(\alpha) - g(\alpha)| + |g(\alpha) - h(\alpha)| \leq \rho(f, g) + \rho(g, h) \). But \( \alpha \in S \) was chosen arbitrarily, so \( \rho(f, h) = \sup_{x \in S} |f(x) - h(x)| \leq \rho(f, g) + \rho(g, h) \).

Example 2.5 is a discrete analogue of Example 2.7, so you should be able to adapt the argument above to that example too.

Definition 2.9. We call the pair \((X, \rho)\) a metric space if \( X \) is a non-empty set and \( \rho \) is a metric on \( X \).

Note that M3 expresses a general triangle inequality for general metric spaces: if \( x, y, z \) are points in \( X \) then the distance from \( x \) to \( z \) cannot exceed the sum of the distances from \( x \) to \( y \) and from \( y \) to \( z \).

Definition 2.10. If \((X, \rho)\) is a metric space and \( A \subset X \) then \( \rho \) is also a metric on \( A \). The metric space \((A, \rho)\) is called a subspace of \((X, \rho)\).

Example 2.11. A continuous function on a bounded closed interval is always bounded. Therefore \( C[a, b] \) is a subspace of \( B[a, b] \) whenever \(-\infty < a < b < +\infty \). (Refer to Examples 2.6 and 2.7. We already noted that the metrics agree.) In fact \( C[a, b] \) is a strict subspace of \( B[a, b] \) since there are bounded functions that are not continuous.

If \( X \) is a linear space, it is often possible to express the metric \( \rho \) in terms of a function of one variable that can be thought of as the length of each element (i.e., its distance from 0).
Definition 2.12. Let \( X \) be a vector space over \( \mathbb{R} \) (or over \( \mathbb{C} \)). A function \( \| \cdot \| : X \to \mathbb{R} \) is called a norm on \( X \) if it satisfies

N1. \( \|x\| > 0 \) if \( x \neq 0 \), and \( \|0\| = 0 \),

N2. \( \|\lambda x\| = |\lambda| \|x\| \) for all \( x \in X \) and \( \lambda \in \mathbb{R} \) (or \( \lambda \in \mathbb{C} \)),

N3. \( \|x + y\| \leq \|x\| + \|y\| \) for all \( x, y \in X \).

A linear space with a norm is called a normed space.

A normed space can be viewed quite naturally as a metric space.

Theorem 2.13. Suppose \( X \) is a linear space with norm \( \| \cdot \| \). The function \( \varrho(x, y) = \|x - y\| \) is a metric on \( X \).

Proof. M1 follows from N1, since \( \varrho(x, x) = \|x - x\| = \|0\| = 0 \), and \( \varrho(x, y) = \|x - y\| > 0 \), when \( x \neq y \). M2 follows from N2, through the chain of equalities

\[
\varrho(x, y) = \|x - y\| = \|-1 \cdot (y - x)\| = |-1| \cdot \|y - x\| = \varrho(y, x).
\]

For M3, observe that

\[
\varrho(x, y) + \varrho(y, z) = \|x - y\| + \|y - z\| \geq \|x - y + y - z\| = \|x - z\| = \varrho(x, z),
\]

where the inequality is N3.

However, not every metric arises in this way; one can have a metric \( \varrho \) on a vector space \( X \) such that \( \varrho(x, 0) \) does not have the properties of a norm.

Example 2.14. In Examples 2.2, 2.3 and 2.4, the metrics are generated by the norms \( \|x\| = |x| \), \( \|x\| = (\sum_{i=1}^{n} |x_i|^2)^{1/2} \) and \( \|x\| = \sum_{i=1}^{n} |x_i| \) respectively.

Example 2.15. The metrics (1) and (2) are generated by the norms

\[
\|f\| = \sup_{x \in [a,b]} |f(x)| \quad \text{and} \quad \|f\| = \sup_{x \in S} |f(x)|.
\]

A further way to obtain metric spaces is by taking a product of other metric spaces. For example, if \( (X_1, \varrho_1) \) and \( (X_2, \varrho_2) \) are both metric spaces, then so is \( (X_1 \times X_2, \varrho) \), where \( \varrho : (X_1 \times X_2)^2 \to \mathbb{R} \) is defined by any of:

- \( \varrho((x_1, x_2), (y_1, y_2)) = \varrho_1(x_1, y_1) + \varrho_2(x_2, y_2) \),

- \( \varrho((x_1, x_2), (y_1, y_2)) = \sqrt{\varrho_1(x_1, y_1)^2 + \varrho_2(x_2, y_2)^2} \), or

- \( \varrho((x_1, x_2), (y_1, y_2)) = \max \{ \varrho_1(x_1, y_1), \varrho_2(x_2, y_2) \} \).

It’s not particularly difficult to prove the above claim, but we won’t do so here.