A brief introduction to metric spaces

David E. Rydeheard*

We describe some of the mathematical concepts relating to metric spaces. This is a brief overview of those topics which are relevant to certain metric semantics of languages.

Definition 1. A *metric space* $(\mathbb{S}, \mu)$ is a set $\mathbb{S}$ together with a binary operation $\mu : \mathbb{S} \times \mathbb{S} \to \mathbf{R}^+$, where $\mathbf{R}^+$ is the set of non-negative real numbers, such that, for all x, y, z in $\mathbb{S}$,

1. $\mu(x, y) = 0 \iff x = y$,
2. $\mu(x, y) = \mu(y, x)$,
3. $\mu(x, z) \leq \mu(x, y) + \mu(y, z)$.

Thus μ is symmetric and satisfies the triangle inequality (the third axiom). If the weaker form of the first axiom

$$\forall x \in \mathbb{S}. \quad \mu(x, x) = 0$$

holds instead then we call μ a pseudometric instead of a metric.

Example 1. Consider the set $\mathbb{S} = \mathbf{R} \times \mathbf{R}$, where $\mathbf{R}$ is the set of real numbers, and the series μ_i for $i \geq 1$:

$$\mu_i((x_1, y_1), (x_2, y_2)) = \sqrt[i]{(|x_1 - x_2|)^i + (|y_1 - y_2|)^i}.$$

Each μ_i is a metric on $\mathbf{R} \times \mathbf{R}$.

For $i = 1$ the metric becomes

$$\mu_1((x_1, y_1), (x_2, y_2)) = |x_1 - x_2| + |y_1 - y_2|$$

and is sometimes called the *city-block* metric.

For $i = 2$, we have the familiar Euclidean metric:

$$\mu_2((x_1, y_1), (x_2, y_2)) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

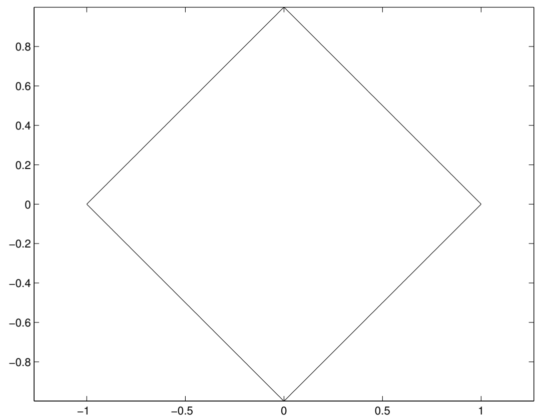
For an example of these two metrics, see Figure 1.

Example 2. Consider bounded functions integrable over an interval $[a, b]$. We may define several simple metrics on the space of these functions:

1. The series of metrics μ_i, $i \geq 0$:

$$\mu_i(f, g) = \sqrt[i]{\int_b^a (|f(x) - g(x)|)^i \, \mathrm{d}x}.$$

*School of Computer Science, University of Manchester, Oxford Road, Manchester M13 9PL, U.K. email: textttdavid.rydeheard@manchester.ac.uk.



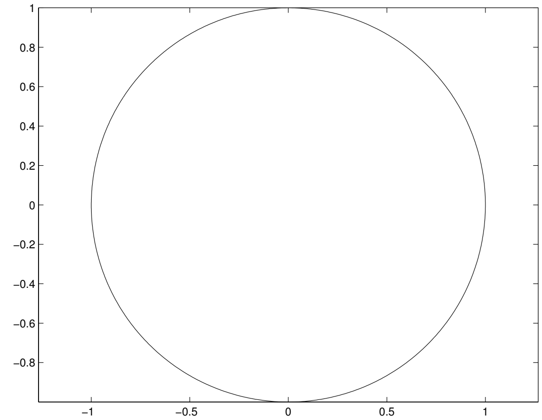


Figure 1: $\{(x, y) : \mu((x, y), (0, 0)) = 1\}$ for (left) the city-block metric and (right) Euclidean metric

2. The *maximum distance metric*, defined as:

$$\mu(f, g) = sup(\{|f(x) - g(x)| : x \in [a, b]\}).$$

The first are the functional analogues of the metrics in Example 1. The second is the limit of the first as i tends to infinity.

Functions into a pseudometric space induce a pseudometric: Given pseudometric space $(\mathbb{S}, \mu)$ and arbitrary function $f : \mathbb{S}' \to \mathbb{S}$, we define a pseudometric on $\mathbb{S}'$ as $\mu'(x, y) = \mu(f(x), f(y))$. For μ a metric and f injective, then μ' is a metric. A common example of this is that any function $f : \mathbb{X} \to \mathbf{R}$ induces a pseudometric μ on $\mathbb{X}$ by $\mu(x, y) = |f(x) - f(y)|$.

Each metric space determines a topological space:

Definition 2. For metric space $(\mathbb{S}, \mu)$, define the topological space $T(\mathbb{S}, \mu)$ to have the set of points $\mathbb{S}$. The open sets are generated from the following base:

$$\{\{y \in \mathbb{S} : \mu(x, y) < d\} : x \in \mathbb{S}, d \in \mathbf{R}^+\}$$

These sets are the open 'spheres', and the open sets $\mathcal{O}(\mathbb{S}, \mu)$ of the space are generated as arbitrary unions of these.

The same construction yields a topological space from a pseudometric. Spaces arising from metrics are Hausdorff. Notice that different metrics may yield the same topology.

There are several useful types of maps between metric spaces:

Definition 3. For metric spaces $(\mathbb{S}, \mu)$ and $(\mathbb{S}', \mu')$, a function $f : \mathbb{S} \to \mathbb{S}'$ is

1. *continuous* if the function induced on the topological spaces $f : T(\mathbb{S}, \mu) \to T(\mathbb{S}', \mu')$ is a continuous map,

2. an *isometry* if $\forall x, y \in \mathbb{S}.\ \mu'(f(x), f(y)) = \mu(x, y)$,

3. a *scaling* if there is a constant $k \in \mathbf{R}$ such that $\forall x, y \in \mathbb{S}.\ \mu'(f(x), f(y)) = k\mu(x, y)$. A bijective scaling is a *similarity*,

4. *partially contractive* if $\forall x, y \in \mathbb{S}.\ \mu'(f(x), f(y)) \leqslant \mu(x, y)$ (*contractive* if the inequality is strict).

We now consider the notion of paths and *shortest paths* (sometimes called *geodesics*) in metric space.

A *path* in a space $(\mathbb{S}, \mu)$ is a continuous function $u : [a, b] \to \mathbb{S}$ where $[a, b]$ ($a \leqslant b$) is the closed interval of reals between a and b. Sometimes conditions stronger than continuity are imposed, but this suffices for the following definitions.

Shortest paths (geodesics) are defined as paths parameterisable by arc-length:

Definition 4. A path $u : [a, b] \to \mathbb{S}$ (with $a \leqslant b$) in a metric space $(\mathbb{S}, \mu)$ is a *shortest path* or *geodesic*, if for all $t, t' \in [a, b]$,

$$\mu(u(t), u(t')) = |t - t'|.$$

This is a limited notion of geodesic applicable to metric spaces which have at least one such geodesic path between every pair of points. The models we consider are of this form. Geodesics may then be unique or there may be a multiplicity (sometimes an infinity) of geodesics between a pair of points.

To calculate geodesic paths, often the metric space is of such a form that the length of a path can be expressed as an integral and then the shortest path is the path (in a certain set of paths) that minimizes the integral, and this can be determined by variational methods in calculus.

Example 3. For the city-block metric in $\mathbf{R}^n$, geodesics are paths which monotonically advance towards the end-points. Thus, a continuous path $u : [0, 1] \to \mathbf{R}^n$ is a geodesic iff for all $s, t \in [0, 1]$, $s \leqslant t \implies u(s)_i \leqslant u(t)_i$ when $u(0)_i \leqslant u(1)_i$ and $u(s)_i \geqslant u(t)_i$ when $u(0)_i \geqslant u(1)_i$, where x_i denotes the i-th projection, $1 \leqslant i \leqslant n$.

For the Euclidean metric in $\mathbf{R}^n$, geodesics are straight lines.

We now turn to subsets of metric spaces:

Definition 5. A *region* of a metric space is a non-empty path-connected subset, i.e. a subset $S \subseteq \mathbb{S}$, such that all pairs of points in S are connected by a path.

Other conditions on regions may be imposed, for example they may be closed and/or bounded. What is the distance between two regions? One possible definition of such a distance is:

Definition 6. The *Hausdorff distance* between two subsets X, Y of metric space $(\mathbb{S}, \mu)$, is given by the expression

$$\mu_H(X, Y) = \max\{\sup_{x \in X} \inf_{y \in Y} (\mu(x, y)), \sup_{y \in Y} \inf_{x \in X} (\mu(x, y))\}$$

and is defined when both sup's and inf's are defined.

On the set of closed, bounded subsets of $\mathbb{S}$, the Hausdorff distance is defined and is a metric.

Example 4. If X and Y are two spheres with radii r and s (with $r \geqslant s$), and distance between centres of d, then the $\mu_H(X, Y) = d + r - s$.

Products of metric spaces: Metric spaces, with isometries as morphisms, do not have cartesian products. However, given metric spaces $(\mathbb{S}, \mu)$ and $(\mathbb{S}', \mu')$, we can define a pseudometric on the product of the spaces $\mathbb{S} \times \mathbb{S}'$, if we are given a function $\nu : \mathbf{R}^+ \times \mathbf{R}^+ \to \mathbf{R}^+$, satisfying the following, for all $u, v, w, x \in \mathbf{R}^+$:

$$\nu(0,0) = 0$$
$$\nu(u + v, w + x) \leqslant \nu(u, w) + \nu(v, x)$$
$$u \leqslant v \text{ and } w \leqslant x \implies \nu(u, w) \leqslant \nu(v, x).$$

The pseudometric $\overline{\mu}$ on $\mathbb{S} \times \mathbb{S}'$ is defined as

$$\overline{\mu}((x, x'), (y, y')) = \nu(\mu(x, y), \mu'(x', y')).$$

If further:

$$\nu(u, v) = 0 \implies u = v = 0$$

then this pseudometric is a metric.

The function ν in the above defines how the two metrics are to be composed. Examples of ν are

1. a linear combination: $\nu(u, v) = \lambda_1 u + \lambda_2 v$ with $\lambda_1, \lambda_2 \geqslant 0$,

2. maximum distance: $\nu(u, v) = \max(u, v)$, and

3. the Euclidean combination: $\nu(u, v) = \sqrt{u^2 + v^2}$.

Products have two roles: allowing us to combine metric spaces modelling different quantities, and also in the definition of metrics for a single quantity based on a combination of separate aspects of the quantity each of which is itself a metric.