Background on Metric Geometry, I

Facundo Mémoli
Math and CSE Departments, OSU.
memoli@math.osu.edu

2014

Table of Contents

Bibliography

- Burago, Burago and Ivanov's book [BBI01].
- Gromov's book [Gro99].
- do Carmo (Differential Geometry of Curves and Surfaces).

Metric space

Definition 1.1. A metric space is a pair (X, d_X) where X is a set and $d_X : X \times X \to \mathbb{R}^+$ with

- $d_X(x, x') = 0$ if and only if $x = x'$.
- $d_X(x, x') = d_X(x', x)$ for all $x, x' \in X$.
- $d_X(x, x'') \leqslant d_X(x, x') + d_X(x', x'')$ for all $x, x', x'' \in X$.

One says that d_X is the metric or distance on X.

If d_X satisfies all but the first condition above, one says that d_X is a *semi-metric* on X.

I will frequently refer to a metric space X with the implicit assumption that its metric is denoted by d_X.

Remark 1.1 (Distance to a set)**.** Let $S \subset X$. We define the distance to S, $d_X(\cdot, S) : X \to \mathbb{R}^+$ by

$$x \mapsto \inf_{s \in S} d_X(x, s).$$

Examples

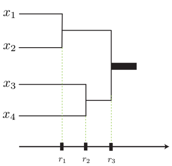
Example 1.1 (Restriction metric). $\Omega \subset \mathbb{R}^n$ and $d_\Omega(\omega, \omega') = \|\omega - \omega'\|$.

Example 1.2 ($\mathbb{S}^n$). Spheres $\mathbb{S}^n$ with "intrinsic" metric. Consider $\mathbb{S}^n \subset \mathbb{R}^{n+1}$ and for $x, x' \in \mathbb{S}^n$ this metric is given by

$$d_{\mathbb{S}^n}(x, x') = 2 \arcsin \frac{\|x - x'\|}{2}$$

Example 1.3 (Ultrametrics). Finite set $\mathbb{X}$ and $u : \mathbb{X} \times \mathbb{X} \to \mathbb{R}^+$ which is symmetric and

$$u(x, x'') \leqslant \max\big(u(x, x'), u(x', x'')\big) \quad \text{for all } x, x', x'' \in \mathbb{X}.$$



$$((u_\theta)) = \begin{array}{c c} & \begin{array}{cccc} x_1 & x_2 & x_3 & x_4 \end{array} \\ \begin{array}{c} x_1 \\ x_2 \\ x_3 \\ x_4 \end{array} & \begin{pmatrix} 0 & r_1 & r_3 & r_3 \\ r_1 & 0 & r_3 & r_3 \\ r_3 & r_3 & 0 & r_2 \\ r_3 & r_3 & r_2 & 0 \end{pmatrix} \end{array}$$

Ultrametrics appear in many applications, including hierarchical clustering (dendrograms).

Compact metric spaces

Definition 1.2. Let X be a metric space and $\varepsilon > 0$. A set $S \subset X$ is called an ε-net for X if $d_X(x, S) \leqslant \varepsilon$ for all $x \in X$.

X is called *totally bounded* if for any $\varepsilon > 0$ there is a finite ε-net for X.

Definition 1.3. For a given $\varepsilon > 0$ a set S in a metric space X is called *ε-separated* if $d_X(s, s') \geqslant \varepsilon$ for all $s, s' \in S$.

Excercise 1. Prove that

1. if there exists a $\frac{\varepsilon}{3}$-net for X with cardinality n, then an ε-separated set in X cannot contain more than n points.

2. A maximal separated ε-set is an ε-net.

Compact metric spaces

An open covering of a topological space is any collection $\mathcal{U}$ of open sets such that $\bigcup_{U\in\mathcal{U}} = X$.

Recall that (by definition) a compact topological space X is one for which any open covering has a finite sub-collection that still covers X.

Also, recall that a *complete metric space* is one for which Cauchy sequences converge. The sequence $\{x_n\}_n \subset X$ is Cauchy, if for any $\varepsilon > 0$ there exists $N \in \mathbb{N}$ s.t. $d_X(x_n, x_m) < \varepsilon$ for all $n, m > N$.

Theorem 1.1. *Let X be a metric space. Then, X is compact if and only if X is complete and totally bounded.*

We denote by $\mathcal{G}$ (for Gromov) the collection of all compact metric spaces.

Hausdorff distance

Let X be a metric space. For a set $S \subset X$ we denote by $B_\varepsilon(S)$ the set of all points x such that $d_X(x,S) < \varepsilon$.

Definition 1.4 (Hausdorff distance)**.** Let $A, B \in 2^X$. The Hausdorff distance between A and B is defined by

$$d_{\mathcal{H}}^X(A,B) := \inf\{\varepsilon > 0 |\, A \subset B_\varepsilon(B) \text{ and } B \subset B_\varepsilon(A)\}.$$

Excercise 2. Prove that

$$d_{\mathcal{H}}^X(A,B) = \max\left(\sup_{a\in A} d_X(a,B), \sup_{b \in B} d_X(b,A)\right).$$

Proposition 1.1. *Let X be a metric space. Then*

1. *$d_{\mathcal{H}}^X(,)$ is a semi-metric on 2^X.*
2. *$d_{\mathcal{H}}^X(A,\overline{A}) = 0$ for any $A \subset X$.*
3. *If $A, B \subset X$ are closed and $d_{\mathcal{H}}^X(A,B) = 0$, then $A = B$.*

Hausdorff distance, cont'd

Let $\mathcal{C}(X)$ denote the collection of all closed subsets of X. Then, we have that $(\mathcal{C}(X), d_{\mathcal{H}}^{X}(,))$ is a metric space. Furthermore, one has

Theorem 1.2 (Blaschke). *If (X, d_X) is compact, then $(\mathcal{C}(X), d_{\mathcal{H}}^{X}(,))$ is also compact.*

That is, we have an application $\mathfrak{H} : \mathcal{G} \to \mathcal{G}$

$$(X, d_X) \mapsto (\mathcal{C}(X), d_{\mathcal{H}}^{X}(,)).$$

Remark 1.2. For a given $S \subset X$, let

$$\mathbf{CovRad}_X\,(S) := \inf\{\varepsilon > 0 |\, X \subseteq B_\varepsilon(S)\}.$$

Clearly, $\mathbf{CovRad}_X\,(S) = d_{\mathcal{H}}^{X}(S, X)$.

Matlab code for Hausdorff distance

Try coding the Hausdorff distance between finite subsets of the plane in Matlab.

Minimal ε-nets in (compact) metric spaces

Lemma 1.1 (Marriage Lemma). *Let Z and Z' be finite sets with a relation $K \subset Z \times Z'$ such that for any $A \subset Z$, $\big|K(A)\big| \geqslant |A|$. Then, there exists a bijection $\varphi : Z \to Z'$ with $(z, \varphi(z)) \in K$ for all $z \in Z$.*

Proposition 1.2. *Fix $\varepsilon > 0$ and assume S, S' are two minimal ε-nets in X (assumed to be compact) with $n = n(\varepsilon)$ points each. Then, there exists a bijection $\varphi : S \to S'$ s.t. $\max_{s \in S} d_X(s, \varphi(s)) \leqslant 2\varepsilon$.*

Proof. Let $R \subset S \times S'$ be given by all those (s, s') s.t. $B_\varepsilon(s) \cap B_\varepsilon(s') \neq \emptyset$. Pick any $A \subset S$ and note that

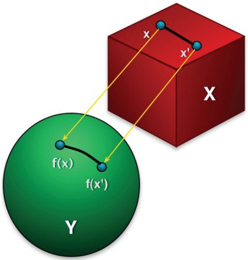
$$R(A) := \bigcup_{a \in A} \{s' \in S' \,|\, (a, s) \in R\}$$

is s.t. $|R(A)| \geqslant |A|$. Otherwise, consider $N := R(A) \cup \big(S \backslash A\big)$. Clearly, N is an ε-net for X and $|N| < n(\varepsilon)$— this contradicts the fact that $n(\varepsilon)$ is the minimal cardinality amongst all ε-nets of X. Then, apply the Marriage Lemma to conclude that there exists a bijection $\varphi : S \to S'$ s.t. $(s, \varphi(s)) \in R$ for all $s \in S$. This means that $B_\varepsilon(s) \cap B_\varepsilon(\varphi(s)) \neq \emptyset$ and hence $d(s, \varphi(s)) \leqslant 2\varepsilon$. $\square$

Maps between metric spaces: Distortion

Definition 1.5 (Distortion)**.** Let (X, d_X) and (Y, d_Y) be metric spaces and $f : X \to Y$ a map. The *distortion* of f is defined as

$$\operatorname{dis}(f) := \sup_{x,x' \in X} \big| d_X(x, x') - d_Y(f(x), f(x')) \big|.$$



We say that a map $f : X \to Y$ is *distance preserving* whenever $\operatorname{dis}(f) = 0$.

An *isometry* between X and Y is any bijective map $f : X \to Y$ which is in addition distance preserving. One says that two metric spaces are *isometric* if there exists an isometry between them.

The set $\operatorname{Iso}(X)$ of all isometries $f : X \to X$ is called the *isometry group* of X.

Isometry: the case of Euclidean sets

If X and Y in the definition of isometry are both $\mathbb{R}^d$, then we are looking at maps $T : \mathbb{R}^d \to \mathbb{R}^d$ such that $\|T(p) - T(q)\| = \|p - q\|$ for all points $p, q \in \mathbb{R}^d$.

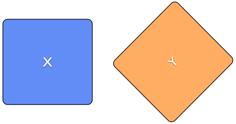
All such T arise as the composition of a translation and a orthogonal transformation. The set of all such maps is denoted $E(d)$ and is called the *Euclidean Group*. If we choose (canonical basis) coordinates on $\mathbb{R}^d$, then any $T \in E(d)$ can be represented as $T(p) = Q\,p + b$ where Q is a $d \times d$ orthogonal matrix (i.e. $Q\,Q^T = I$) and b is (translation) vector.

If Q is an orthogonal matrix, then $\|Q\,p\| = \|p\|$ for all points p. This means that whenever $T \in E(d)$, then $< T(p), T(q) >=< p, q >$ for all points p, q.

Isometry: the case of finite Euclidean sets

Lemma 1.2 (Folklore Lemma). *Let $A = \{a_1, a_2, \ldots, a_n\}$ and $B = \{b_1, b_2, \ldots, b_n\}$ two sets in $\mathbb{R}^d$ such that $\|a_i - a_j\| = \|b_i - b_j\|$ for all $i, j = 1, \ldots, n$. Then, there exists an isometry of ambient space $T : \mathbb{R}^d \to \mathbb{R}^d$ such that $b_i = T(a_i)$ for each i.*

This is actually quite interesting: from purely *intrinsic* information (the interpoint distances of points in A and B we can deduce the existence of an *extrinsic* object (the transformation T) with strong properties. This happens because we are assuming that both A and B are "special", namely they live in Euclidean space.



Excercise 3. Try to prove this theorem for youself. Or try to find and study the proof.

Isometric embedding

Definition 1.6. A map $f : X \to Y$ between metric spaces X and Y is an *isometric embedding* of X into Y if $f(X) \subset Y$ endowed with the restriction of the metric from Y is isometric to X.

Example 1.4. $\mathbb{S}^2$ with the intrinsic metric does not admit an isometric embedding into any $\mathbb{R}^k$, $k \in \mathbb{N}$.

Excercise 4 (Kuratowski's embedding)**.** Let X be a compact metric space. Consider $\big(C(X), \|\cdot\|_{L^\infty}\big)$, the metric space of all real valued continuous functions on X, where the metric is the L^∞ norm. Attach to each $x \in X$ the function $f_x := d_X(\cdot, x) : X \to \mathbb{R}^+$. Then,

$$\|f_x - f_{x'}\|_\infty = d_X(x, x') \quad \text{for all } x, x' \in X.$$

Lipschitz maps and dilatation

Definition 1.7. A map $f : X \to Y$ between metric spaces X and Y is called *Lipschitz* if there exists $L \geqslant 0$ s.t.

$$d_Y(f(x), f(x')) \leqslant L \cdot d_X(x, x') \quad \text{for all } x, x' \in X.$$

Any such L is called a *Lipschitz constant* for f. The minimal Lipschitz constant of a map f is called the *dilatation of f* and denoted $\mathrm{dil}(f)$.

A map with Lipschitz constant 1 is called *non-expanding*.

Distance preserving maps in compact spaces

A compact metric space cannot be isometric to a proper subset of itself.

Theorem 1.3. *Let X be a compact metric space and $f : X \to X$ be a distance preserving map. Then $f(X) = X$.*

Proof. Assume $p \in X \backslash f(X)$. Since $f(X)$ is compact (and hence closed) there exists $\varepsilon > 0$ s.t. $B_\varepsilon(p) \cap f(X) = \emptyset$. Let n be the maximal cardinality of an ε separated set in X and $S \subset X$ be an ε-separated set with cardinality n. Then, $f(S)$ is also ε-separated. Also,

$$d_X(p, f(S)) \geqslant d_X(p, f(X)) \geqslant \varepsilon$$

and therefore $f(S) \cup \{p\}$ is also ε-separated but with cardinality $n+1$, which contradicts the maximality of n. $\square$

Non-expanding maps in compact spaces

Theorem 1.4. *Let X be a compact metric space. Then,*

1. *Any non-expanding* surjective *map is an isometry.*
2. *If a map $f : X \to X$ is non-contracting: $d_X(f(x), f(x')) \geqslant d_X(x, x')$ for all $x, x' \in X$, then f is an isometry.*

Proof. We prove (1). Assume p, q are such that $d_X(f(p), f(q)) < d_X(p, q)$ for some $p, q \in X$. Fix such a pair of points and pick $\varepsilon > 0$ s.t. $d_X(f(p), f(q)) < d_X(p, q) - 5\varepsilon$. Let n be such that there exists at least one ε-net in X of cardinality n. Let $\mathcal{N}_n \subset X^n$ the collection of all n-tuples of points that form ε-nets in X. This set is closed in X and hence compact. Define the function $D : X^n \to \mathbb{R}^+$ given by

$$(x_1, \ldots, x_n) \mapsto \sum_{i,j=1}^{n} d_X(x_i, x_j).$$

This function is continuous and therefore attains a minimum on $\mathcal{N}_n$. Let $S = (x_1, \ldots, x_n) \in \mathcal{N}_n$ be a minimizer. Since f is non-expanding and surjective then $f(S) \in \mathcal{N}_n$. Also, $D(f(S)) \leqslant D(S)$ and hence $d_X(x_i, x_j) = d_X(f(x_i), f(x_j))$ for all $i, j \in \{1, \ldots, n\}$.

Let i_0 and j_0 be s.t. $d_X(p, x_{i_0}), d_X(q, x_{j_0}) \leqslant \varepsilon$. Then, one has $d_X(x_{i_0}, x_{j_0}) \geqslant d_X(p, q) - 2\varepsilon$ and $d_X(f(x_{i_0}), f(x_{j_0})) \leqslant d_X(f(p), f(q)) + 2\varepsilon \leqslant d_X(p, q) - 3\varepsilon$. This gives $d_X(x_{i_0}, x_{j_0}) > d_X(f(x_{i_0}), f(x_{j_0}))$, a contradiction. ☐

Diameter, inradius, etcetera

Let $X \in \mathcal{G}$. Define

- **separation**: $X \mapsto \mathbf{sep}\,(X) := \min\{d_X(x,x'),\, x \neq x'\}$.
- **diameter**: $X \mapsto \mathbf{diam}\,(X) := \max_{x,x'} d_X(x,x')$.
- **inradius**: $X \mapsto \mathbf{rad}\,(X) := \min_x \max_{x'} d_X(x,x')$.
- **eccentricity**: $X \mapsto \mathbf{ecc}_X : X \to \mathbb{R}^+$. It is given by $\mathbf{ecc}_X(x) = \max_{x'} d_X(x,x')$, $x \in X$.
- **curvature sets**: $(X,k) \mapsto K_k(X) \subset \left(\mathbb{R}^+\right)^{k\times k}$, the collection of all $k \times k$ symmetric matrices $((d_X(x_i,x_j)))$ where $(x_1,x_2,\ldots,x_k) \in X^k$.
- Packing and covering numbers and others: consider things like
  - $\mathbf{xt}_k(X) := \left(C_2^k\right)^{-1} \max\{\sum_{i>j}^k d_X(x_i,x_j), (x_1,\ldots,x_k) \in X^k\}$,
  - $\mathbf{cov}_k(X) := \min\{\varepsilon \geqslant 0$ s.t. exists ε-net S for X, with $|S| = k\}$
  - $\mathbf{cap}_k(X) := \max\{\varepsilon \geqslant 0$ s.t. exists S with $|S| = k$ and $\mathbf{sep}\,(S) \geqslant \varepsilon\}$

Question 1.1. What happens to these invariants if I "perturb" X slightly? How can one define a notion of perturbation of metric spaces?

some more invariants

- **CapNbr**$(X, \varepsilon) := \max\left\{|S|; S \subset X \text{ with } B_{\varepsilon/2}(x_i) \cap B_{\varepsilon/2}(x_j) = \emptyset,\ i \neq j\right\}$
- **CovNbr**$(X, \varepsilon) := \min\{|S|; S \subset X \text{ with } X \subseteq B_\varepsilon(S)\}$

Excercise 5. Prove that $\mathbf{CapNbr}(X, \varepsilon) \geqslant \mathbf{CovNbr}(X, \varepsilon)$. for all $\varepsilon \geqslant 0$.

ε-Isometries

We seek a *relaxation* of the notion of isometries. We need to preserve distances alright, but we also need to make sure we fill in the target space with the image of the source via the "approximate isometry map". This suggests:

Definition 1.8. One says that a map $f : X \to Y$ is a ε-isometry between metric spaces X and Y if

- $\mathrm{dis}(f) \leqslant \varepsilon$ and
- $f(X)$ is an ε-net for Y.

Note that we do not require f to be continuous.

Excercise 6. Prove that a 0-isometry coincides with an isometry in the usual sense.

Length structures

Consider a metric space (X, d_X). We'll construct a new metric d_X^* over X. For each continuous path $\gamma : [0,1] \to X$ we consider its *length*

$$L_{d_X}(\gamma) := \sup\{\sum_{i=1}^{N} d_X(\gamma(t_i), \gamma(t_{i+1}))\}$$

where the supremum is taken over all *partitions* of $[0,1]$. A set of points $P = \{t_1, t_2, \ldots, t_N\}$ is a partition of $[0,1]$ if $0 = t_1 \leqslant t_2 \leqslant \cdots \leqslant t_N = 1$. We say that a curve γ is *rectifiable* whenver its length is finite.

The intrinsic metric d_X^* on X is defined as follows: for each pair $x, x' \in X$ we consider $\Gamma_X(x, x')$ the set of all continuos paths joining x to x'. Then, we define

$$d_X^*(x, x') := \inf\{L_{d_X}(\gamma); \gamma \in \Gamma_X(x, x')\}.$$

Notice that if the points x and x' cannot be connected by a continous curve, then the definition above does not make sense. So when that is the case, one would (informally) say that $d_X^*(x, x') = \infty$. Notice that this can be the case even if $d_X(x, x')$ is finite.

Geodesic metric spaces

Consider a metric space (X, d_X). We'll construct a new metric d_X^* over X. For each continuous path $\gamma : [0,1] \to X$ we consider its *length*

$$L_{d_X}(\gamma) := \sup\{\sum_{i=1}^{N} d_X(\gamma(t_i), \gamma(t_{i+1}))\}$$

where the supremum is taken over all *partitions* of $[0,1]$. A set of points $P = \{t_1, t_2, \ldots, t_N\}$ is a partition of $[0,1]$ if $0 = t_1 \leqslant t_2 \leqslant \cdots \leqslant t_N = 1$. We say that a curve γ is *rectifiable* whenver its length is finite.

The intrinsic metric d_X^* on X is defined as follows: for each pair $x, x' \in X$ we consider $\Gamma_X(x,x')$ the set of all continuos paths joining x to x'. Then, we define

$$d_X^*(x,x') := \inf\{L_{d_X}(\gamma);\, \gamma \in \Gamma_X(x,x')\}.$$

Notice that if the points x and x' cannot be connected by a continous curve, then the definition above does not make sense. So when that is the case, one would (informally) say that $d_X^*(x,x') = \infty$. Notice that this can be the case even if $d_X(x,x')$ is finite.

If $d_X^* = d_X$, then one says that d_X is *intrinsic* and that (X, d_X) is a *path metric space* or also sometimes *length space*. One says that (X, d_X) is *geodesic* if for any pair of points x, x' there exists $\gamma \in \Gamma_X(x,x')$ such that $L_{d_X}(\gamma) = d_X(x,x')$.

A zoo of metric spaces

- We saw examples of different metrics on spheres.
- We saw ultrametric spaces.
- More general are *trees*.
- Trees are a special class of *graphs*, which can also be regarded as metric spaces.
- Graphs are frequently endowed with the *path length metric*.
- Path length metric spaces are a subclass of metric spaces.
- Riemannian manifolds are a subclass of Length spaces. The metric on Riemannian manifolds is usually referred to as the *geodesic metric*.

Bibliography

References

[BBI01] D. Burago, Y. Burago, and S. Ivanov. *A Course in Metric Geometry*, volume 33 of *AMS Graduate Studies in Math.* American Mathematical Society, 2001.

[Gro99] Misha Gromov. *Metric structures for Riemannian and non-Riemannian spaces*, volume 152 of *Progress in Mathematics*. Birkhäuser Boston Inc., Boston, MA, 1999.