HAUSDORFF DISTANCE AND MEASURABILITY

Contents

In this class, we will tackle reconstruction problems. Namely, given random variables $X_1, \ldots, X_n$, we build — with various methods depending on the context — estimators $\hat{K} = \hat{K}(X_1, \ldots, X_n)$ aiming at approximate a target compact subset $K \subset \mathbb{R}^D$. Hence, we have to make clear what "approximate" means for compact sets. For this, we use the Hausdorff distance $\mathrm{d_H}$. Consequently, we have to clarify what to be a compact sets-valued estimator means, or equivalently, describe measurability properties in the space of compact subsets endowed with the Hausdorff distance. To ensure not to focus on technical details about measurability later on, we choose to address them in this note.

Roughly speaking, the take-away message is that the class of compact subsets of a metric space behaves as well as the metric space itself. Hence, random variables with values in it do so.

1. Hausdorff Distance

Let $(\mathcal{D}, \mathrm{d})$ be a metric space. This class will only tackle the case $(\mathcal{D}, \mathrm{d}) = (\mathbb{R}^D, \|\cdot\|)$. However, we state Hausdorff distance properties in full generality to emphasize the key points that have our case work.

We let $\mathcal{K}(\mathcal{D})$ denote the set of nonempty compact subsets of $(\mathcal{D}, \mathrm{d})$. For $x \in \mathcal{D}$ and $K \subset \mathcal{D}$, the distance from x to K is

$$\mathrm{d}_K(x) = \inf \left\{ \mathrm{d}(x, y), y \in K \right\}.$$

One easily checks that $\mathrm{d}_K(.)$ is a 1-Lipschitz map. Let us define the Hausdorff distance.

Definition 1.1. For two compact subsets $A, B \subset \mathbb{R}^D$, the *Hausdorff distance* between A and B is defined by

$$\mathrm{d_H}(A, B) = \max\Big\{ \sup_{a \in A} \mathrm{d}_B(a), \sup_{b \in B} \mathrm{d}_A(b) \Big\}.$$

$\mathrm{d_H}$ is a distance on the space $\mathcal{K}(\mathcal{D})$ of nonempty compact subsets of $(\mathcal{D}, \mathrm{d})$.

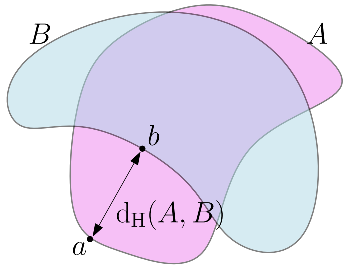

FIGURE 1. The Hausdorff distance between two subsets A and B of the plane. In this example, $\mathrm{d_H}(A,B)$ is the distance between the point a in A which is the farthest from B and its nearest neighbor b on B.

Proof. It is clear from the definition that $\mathrm{d_H}(\cdot,\cdot)$ is finite on compact sets, and symmetric. Moreover, if $\mathrm{d_H}(A,B)=0$, then for all $a\in A$, $\mathrm{d}_B(a)=0$. Since B is a closed subset, we get $A\subset B$. Symmetrically, we get $B\subset A$, which shows that $\mathrm{d_H}$ is separated. Let now $A,B,C\in\mathcal{K}(\mathcal{D})$. Since $\mathrm{d}_B(\cdot)$ is 1-Lipschitz, for all $a\in A$ and $c\in C$, $\mathrm{d}_B(a)\leqslant\mathrm{d}_B(c)+\mathrm{d}(a,c)$. By definition, $\mathrm{d}_B(c)\leqslant\mathrm{d_H}(C,B)$. Hence,

$$\begin{aligned}\mathrm{d}_B(a)&\leqslant\mathrm{d_H}(C,B)+\inf_{c\in C}\mathrm{d}(a,c)\\&=\mathrm{d_H}(C,B)+\mathrm{d}_C(a)\\&\leqslant\mathrm{d_H}(C,B)+\mathrm{d_H}(A,C),\end{aligned}$$

so that $\sup_{a\in A}\mathrm{d}_B(a)\leqslant\mathrm{d_H}(C,B)+\mathrm{d_H}(A,C)$. By a symmetric argument, we get $\sup_{b\in B}\mathrm{d}_A(b)\leqslant\mathrm{d_H}(C,A)+\mathrm{d_H}(B,C)$, which gives the triangle inequality $\mathrm{d_H}(A,B)\leqslant\mathrm{d_H}(A,C)+\mathrm{d_H}(C,B)$. □

An equivalent formulation of $\mathrm{d_H}$ can be written in terms of offsets. Recall that the *r-offset* of K is defined as

$$K^r=\{x\in\mathcal{D},\mathrm{d}_K(x)\leqslant r\}\,,$$

that is, the set of ambient points that are at distance less than or equal to r from K.

PROPOSITION 1.2. *For all $A,B\in\mathcal{K}(\mathcal{D})$,*

$$\mathrm{d_H}(A,B)=\inf\{r>0,A^r\supset B\text{ and }B^r\supset A\}\,.$$

Proof of Proposition 1.2. By definition, for all $a\in A$, $\mathrm{d}_B(a)\leqslant\mathrm{d_H}(A,B)$, which yields $B^{\mathrm{d_H}(A,B)}\supset A$. Symmetrically, $A^{\mathrm{d_H}(A,B)}\supset B$, and hence

$$\mathrm{d_H}(A,B)\geqslant\inf\{r>0,A^r\supset B\text{ and }B^r\supset A\}\,.$$

Conversely, without loss of generality, there exists a point $a_0\in A$ such that $\mathrm{d}_B(a_0)=\mathrm{d_H}(A,B)$. Hence, for all $r<\mathrm{d_H}(A,B)$, $a_0\notin B^r$ and in particular $B^r\not\supset A$. Hence the result. □

2. THE METRIC SPACE $(\mathcal{K}(\mathcal{D}),\mathrm{d_H})$

The Hausdorff distance is a rigid distance, in the sense a single point added to a set — say, an outlier — can have the Hausdorff distance blow up, since $\mathrm{d_H}(A,A\cup x)=\mathrm{d}_A(x)$. It plays the role of a L^∞ dissimilarity in the space of compact sets. One can make this idea precise by identifying a compact subset $A \subset \mathcal{D}$ to its distance function $\mathrm{d}_A(\cdot)$, which is locally bounded.

PROPOSITION 2.1. *The map*

$$\begin{aligned}(\mathcal{K}(\mathcal{D}), \mathrm{d}_{\mathrm{H}}) &\longrightarrow (\mathcal{C}(\mathcal{D}, \mathbb{R}_+), \|\cdot\|_\infty) \\ A &\longmapsto \mathrm{d}_A(\cdot)\end{aligned}$$

is an isometry. In other words, for all $A, B \in \mathcal{K}(\mathcal{D})$,

$$\mathrm{d}_{\mathrm{H}}(A, B) = \sup_{x \in \mathcal{D}} |\mathrm{d}_A(x) - \mathrm{d}_B(x)|.$$

Moreover, if $(\mathcal{D}, \mathrm{d})$ *is complete,* $(\mathcal{K}(\mathcal{D}), \mathrm{d}_{\mathrm{H}})$ *is closed in* $(\mathcal{C}(\mathcal{D}, \mathbb{R}_+), \|\cdot\|_\infty)$.

Proof. For all $x \in A$, $\mathrm{d}_A(x) = 0$, so that

$$\sup_{a \in A} \mathrm{d}_B(a) = \sup_{x \in A} \mathrm{d}_B(x) - \mathrm{d}_A(x) \leqslant \sup_{x \in \mathcal{D}} \mathrm{d}_B(x) - \mathrm{d}_A(x).$$

We now prove the reverse inequality. For $x \in \mathcal{D}$, write $\pi_A(x)$ for any element of A such that $\mathrm{d}_A(x) = \mathrm{d}(x, \pi_A(x))$. Then, since $\mathrm{d}_B(\cdot)$ is 1-Lipschitz,

$$\begin{aligned}\mathrm{d}_B(x) - \mathrm{d}_A(x) &= \mathrm{d}_B(x) - \mathrm{d}(x, \pi_A(x)) \\ &\leqslant \mathrm{d}_B(\pi_A(x)) \\ &\leqslant \sup_{a \in A} \mathrm{d}_B(a),\end{aligned}$$

which yields the desired reverse bound, and hence

$$\sup_{a \in A} \mathrm{d}_B(a) = \sup_{x \in \mathcal{D}} \mathrm{d}_B(x) - \mathrm{d}_A(x).$$

By symmetry,

$$\sup_{b \in B} \mathrm{d}_A(b) = \sup_{x \in \mathcal{D}} \mathrm{d}_A(x) - \mathrm{d}_B(x).$$

Conclude writing

$$\begin{aligned}\sup_{x \in \mathcal{D}} |\mathrm{d}_A(x) - \mathrm{d}_B(x)| &= \max\left\{\sup_{x \in \mathcal{D}} \mathrm{d}_B(x) - \mathrm{d}_A(x), \sup_{x \in \mathcal{D}} \mathrm{d}_A(x) - \mathrm{d}_B(x)\right\} \\ &= \max\left\{\sup_{a \in A} \mathrm{d}_B(a), \sup_{b \in B} \mathrm{d}_A(b)\right\} \\ &= \mathrm{d}_{\mathrm{H}}(A, B).\end{aligned}$$

Finally, if $(\mathcal{D}, \mathrm{d})$ is complete, the closedness of $\mathcal{K}(\mathcal{D})$ in the space of continuous functions is proved in Lemma 3.1.1 of [2]. □

Remark 2.2. Actually, we proved the (slightly) more precise identity

$$\mathrm{d}_{\mathrm{H}}(A, B) = \sup_{x \in K} |\mathrm{d}_A(x) - \mathrm{d}_B(x)|,$$

for all $A \cup B \subset K \subset \mathcal{D}$, meaning that one can restrict the distance functions to the domain $A \cup B$ to compare A and B. When measuring the dissimilarity between compact subsets, we can somehow restrict to the geometry of $(A \cup B, d)$.

By identifying a compact subset with its associated distance function, one can see $(\mathcal{K}(\mathcal{D}), \mathrm{d_H})$ as a closed subset of $(\mathcal{C}(\mathcal{D}, \mathbb{R}_+), \|\cdot\|_\infty)$. Consequently, it inherits its usual topological and metric properties. Conversely, one has the isometric closed inclusion

$$\begin{aligned}(\mathcal{D}, \mathrm{d}) &\longrightarrow (\mathcal{K}(\mathcal{D}), \mathrm{d_H})\\ x &\longmapsto \{x\}\,,\end{aligned}$$

that allows to identify a point x to the singleton $\{x\}$. Hence, roughly speaking, $(\mathcal{K}(\mathcal{D}), \mathrm{d_H})$ cannot have better metric properties than $(\mathcal{D}, \mathrm{d})$. We recall that a metric space is said to be boundedly compact if all its closed bounded subsets are compact. In particular, a boundedly compact metric space is complete.

Proposition 2.3. *Let $(\mathcal{D}, \mathrm{d})$ be a metric space.*

(i) $(\mathcal{D}, \mathrm{d})$ *is separable if and only if* $(\mathcal{K}(\mathcal{D}), \mathrm{d_H})$ *is separable,*
(ii) $(\mathcal{D}, \mathrm{d})$ *is compact if and only if* $(\mathcal{K}(\mathcal{D}), \mathrm{d_H})$ *is compact,*
(iii) $(\mathcal{D}, \mathrm{d})$ *is boundedly compact if and only if* $(\mathcal{K}(\mathcal{D}), \mathrm{d_H})$ *is boundedly compact,*
(iv) $(\mathcal{D}, \mathrm{d})$ *is complete if and only if* $(\mathcal{K}(\mathcal{D}), \mathrm{d_H})$ *is complete,*
(v) $(\mathcal{D}, \mathrm{d})$ *is Polish if and only if* $(\mathcal{K}(\mathcal{D}), \mathrm{d_H})$ *is Polish.*

Proof. (i) For the direct sense, notice that a dense sequence $\{x_i\}_{i\in\mathbb{N}}$ of $\mathcal{D}$ provides the countable family $\{\cup_{i\in I}\{x_i\}\}_{\text{finite } I\subset\mathbb{N}}$ which is dense in $\mathcal{K}(\mathcal{D})$. Conversely, if $\mathcal{K}(\mathcal{D})$ is separable, so is $\mathcal{D}$, as a closed subset of the metric space $\mathcal{K}(\mathcal{D})$.
(ii) If $(\mathcal{D}, \mathrm{d})$ is compact, then $\mathcal{K}(\mathcal{D}) \simeq \{\mathrm{d}_A(\cdot)\}_{A\in\mathcal{K}(\mathcal{D})}$ is an equicontinuous and relatively compact family of $\mathcal{C}(\mathcal{D}, \mathbb{R}_+)$, with $\mathcal{D}$ compact. Hence, it is compact. Conversely, if $\mathcal{K}(\mathcal{D})$ is compact, so is $\mathcal{D}$, as a closed subset of $\mathcal{K}(\mathcal{D})$.
(iii) Follows from the same argument as (ii) by localizing.
(iv) From Proposition 2.1, $\mathcal{K}(\mathcal{D})$ is a closed subset of the complete space space $\mathcal{C}(\mathcal{D}, \mathbb{R}_+)$, so it is complete. Conversely, if $\mathcal{K}(\mathcal{D})$ is complete, so is $\mathcal{D}$, as a closed subset of $\mathcal{K}(\mathcal{D})$.
(v) Is a rephrasing of (i) with (iv).

□

To avoid measure-theoretic difficulties, the mildest framework commonly adopted to develop probability theory is random variables with values in Polish spaces [4]. Hence, working in $(\mathcal{K}(\mathcal{D}), \mathrm{d_H})$ when $(\mathcal{D}, \mathrm{d})$ is Polish will have all the usual probability theory tools operate in a non-pathological way. In particular, manipulating random variables in $(\mathcal{K}(\mathbb{R}^\mathcal{D}), \mathrm{d_H})$ will not raise any specific issue.

3. Compact Set-Valued Random Variables

Now that we made sure handling compact sets-valued random variables is not problematic, let us describe some of them in the case $(\mathcal{D}, \mathrm{d}) = (\mathbb{R}^D, \|\cdot\|)$. We give a few examples of measurable maps in $\mathcal{K}(\mathbb{R}^D)$ endowed with the Borel σ-field associated to the Hausdorff metric $\mathrm{d_H}$. We let $\mathcal{C}(\mathbb{R}^D, \mathbb{R}^D)$ denote the set of continuous map from $\mathbb{R}^D$ to itself, that we endow with the topology of the uniform convergence on compact sets, and its Borel σ-field.

Proposition 3.1. *Equip $\mathcal{K}(\mathbb{R}^D)$ with the Borel σ-field associated to the Hausdorff metric $\mathrm{d_H}$. Then the following maps are measurable:*

(i) $\mathbb{R}^D \ni x \longmapsto \{x\}$, *for all* $x \in \mathbb{R}^D$,
(ii) $\mathcal{C}(\mathbb{R}^D, \mathbb{R}^D) \times \mathcal{K}(\mathbb{R}^D) \ni (f, A) \mapsto f(A) = \big\{f(x), x \in A\big\}$,
(iii) $\mathcal{K}(\mathbb{R}^D) \times \mathcal{K}(\mathbb{R}^D) \ni (A, B) \longmapsto A \cup B$,
(iv) $\mathcal{K}(\mathbb{R}^D) \ni A \mapsto \mathrm{conv}(A)$.

Proof. We actually prove continuity of the considered map, which is stronger than measurability.

(i) It is an isometry.
(ii) To prove that the map $(f, A) \mapsto f(A)$ is jointly measurable, it is sufficient to prove continuous in each variable separately, from Lemma 4.51 in [1]. Fix $A \in \mathcal{K}(\mathbb{R}^D)$. Then for all f, g continous, $\mathrm{d_H}(f(A), g(A)) \leqslant \sup_{x\in A} |f(x) - g(x)|$, which goes to zero when g converges to f on the compact A. Let now f be fixed. Then for all $A \in \mathcal{K}(\mathbb{R}^D)$, consider $K = A^1$, the offset of radius 1 of A. Then f is uniformly continuous on the compact set K. Hence, for all $\varepsilon > 0$, there exists $\eta > 0$ such that for all $x, y \in K$ such that $\|y - x\| \leqslant \eta$, $\|f(y) - f(x)\| \leqslant \varepsilon$. Hence, for $\mathrm{d_H}(A, B) \leqslant \eta \wedge 1$, we get $\mathrm{d_H}(f(A), f(B)) \leqslant \varepsilon$, which proves continuity of $B \mapsto f(B)$ at A, and concludes the proof.
(iii) Writting $r = \max\big\{\mathrm{d_H}(A_1, A_2), \mathrm{d_H}(B_1, B_2)\big\}$, we have $(A_1 \cup B_1)^r = A_1^r \cup B_1^r \supset A_2 \cup B_2$. Symmetrically, $(A_2 \cup B_2)^r \supset A_1 \cup A_2$, so that
$$\mathrm{d_H}(A_1 \cup B_1, A_2 \cup B_2) \leqslant \max\big\{\mathrm{d_H}(A_1, A_2), \mathrm{d_H}(B_1, B_2)\big\}.$$
(iv) For any convex combination $\overline{a} = \sum_i \lambda_i a_i \in \mathrm{conv}(A)$ of elements of A, considering convex combinations $\overline{b} = \sum_i \lambda_i b_i$ for $b_i \in B$ clearly yields
$$\mathrm{d}_{\mathrm{conv}(B)}(\overline{a}) \leqslant \sum_i \lambda_i \mathrm{d}_B(a_i) \leqslant \sum_i \lambda_i\, \mathrm{d_H}(A, B) = \mathrm{d_H}(A, B).$$
Symmetrically, we obtain $\mathrm{d}_{\mathrm{conv}(A)}(\bar{b})$ for all $\bar{b} \in \mathrm{conv}(B)$. Hence,
$$\mathrm{d_H}(\mathrm{conv}(A), \mathrm{conv}(B)) \leqslant \mathrm{d_H}(A, B).$$
□

By composition, Proposition 3.1 actually allows to describe a wide variety of measurable maps. For instance, a simplicial complex is a finite union of simplicies, and simplicies are convex hulls of finite sets. As a consequence, (i),(ii) and (iv) show that simplicial complexes $\hat{M}$ built on top of a random point cloud $\mathbb{X}_n$ for which the presence of each simplex is determined by a measurable event, yield estimators. Similarly, the union of local polynomial patches are measurable from (ii) and (iii).

4. Further Sources

For (much) more details about measurability in classes of subspaces, we refer to [3], and to [2] for the functional approach we adopted.

References

[1] Charalambos D. Aliprantis and Kim C. Border. *Infinite dimensional analysis.* Springer, Berlin, third edition, 2006. A hitchhiker's guide.
[2] Gerald Beer. *Topologies on closed and closed convex sets*, volume 268 of *Mathematics and its Applications.* Kluwer Academic Publishers Group, Dordrecht, 1993.

[3] Georges Matheron. *Random sets and integral geometry.* John Wiley & Sons, New York-London-Sydney, 1975. With a foreword by Geoffrey S. Watson, Wiley Series in Probability and Mathematical Statistics.
[4] Kalyanapuram R. Parthasarathy. *Probability measures on metric spaces.* AMS Chelsea Publishing, Providence, RI, 2005. Reprint of the 1967 original.