Fréchet-Stable Signatures Using Persistent Homology

Don Sheehy
Problem: Compare Shapes. (want a metric)

Example: GPS traces.

Idea: Just run Ripser. What do you expect to get? How much does the sampling density matter?
Our target: **Hausdorff Distance**

Let $A, B$ be compact subsets of a metric space.

The *directed Hausdorff distance* is

$$d^+_H(A, B) := \max_{a \in A} \min_{b \in B} d(a, b)$$

The *Hausdorff distance* is

$$d_H(A, B) := \max \left\{ d^+_H(A, B), d^+_H(B, A) \right\}$$

*at first*
Our target: Hausdorff Distance

Let $A, B$ be compact subsets of a metric space.

The **directed Hausdorff distance** is

$$d_{H^+}(A,B) := \max_{a \in A} \min_{b \in B} d(a,b)$$

The **Hausdorff distance** is

$$d_H(A,B) := \max \{ d_{H^+}(A,B), d_{H^+}(B,A) \}$$

Equivalently, let $A^r := \bigcup_{a \in A} \text{ball}(a,r)$ (and similarly for $B^r$).

Then,

$$d_H := \min \{ r : A \subseteq B^r \text{ and } B \subseteq A^r \}$$

*Interleaving!*

* at first
Hausdorff Stability

$A, B$ compact subsets of a metric space

\[ d_B(\text{Pers}(\text{VR}(A)), \text{Pers}(\text{VR}(B))) \leq d_H(A, B) \]

Also true for offsets:

\[ A = (A^*)_{\alpha \geq 0}, \quad B = (B^*)_{\alpha \geq 0} \]

\[ d_B(\text{Pers}(A), \text{Pers}(B)) \leq d_H(A, B) \]
Trajectories: \( f, g : [0, 1] \to \mathbb{R}^2 \)

Images: \( A = \text{im } f \quad B = \text{im } g \)

Samples: \( \hat{A} \subset A \quad \hat{B} \subset B \quad d_H(\hat{A}, \hat{A}) \leq \varepsilon, \quad d_H(\hat{B}, \hat{B}) \leq \varepsilon \) \quad (\star)

Persistence modules: \( R_X := \text{Pers}(\text{VR}(X)) \quad \text{for } X \in \{A, B, \hat{A}, \hat{B}\} \)
Trajectories: $f, g: [0, 1] \to \mathbb{R}^2$

Images: $A = \text{im } f, B = \text{im } g$

Samples: $\hat{A} \subseteq A, \hat{B} \subseteq B,$ $d_H(A, \hat{A}) \leq \varepsilon, d_H(B, \hat{B}) \leq \varepsilon$  (*)

Persistence modules: $R_{X} := \text{Pers} (VR(X))$ for $X \in \{A, B, \hat{A}, \hat{B}\}$

$$d_B(R_{\hat{A}}, R_{\hat{B}}) \leq d_H(\hat{A}, \hat{B}) \quad \text{(by stability)}$$

$$\leq d_H(\hat{A}, A) + d_H(A, B) + d_H(B, \hat{B}) \quad \text{(by } \Delta\text{-inequality)}$$

$$\leq d_H(A, B) + 2\varepsilon \quad \text{(by (\ast))}$$

Is this good?
The function Perspective

Let $f: X \to \mathbb{R}$, where $X$ is some top. space.

Sets: $F^\alpha := f^{-1}([-\infty, \alpha]) = \{ x \in X : f(x) \leq \alpha \}$

Filtration: $F := (F^\alpha)$

* notation: $\text{Pers}(f) := \text{Pers}(F)$

** Caveat: tameness.  "Persistence of a function"
Hausdorff Again

Let $A, B$ be compact subsets of a metric space $X$

Distance functions: $r_A, r_B : X \rightarrow \mathbb{R}$

\[
\begin{align*}
  r_A(x) &:= \min_{a \in A} d(a, x) \\
  r_B(x) &:= \min_{b \in B} d(b, x)
\end{align*}
\]

$A^r := r_A^{-1}([0, r])$, $B^r := r_B^{-1}([0, r])$

Fact: $d_H(A, B) = \|r_A - r_B\|_\infty = \max_{x \in X} |r_A(x) - r_B(x)|$

Stability for tame functions: $d_B(\text{Pers}(f), \text{Pers}(g)) \leq \|f - g\|_\infty$

\[\Rightarrow\text{Hausdorff Stability}\]
Persistence Equivalence

Take the homology of a filtration over a field.
Persistence Equivalence

Take the homology of a filtration over a field

\[ H_*(F^x) \xrightarrow{F^x \rightarrow G^x} H_*(G^x) \]

\[ H_*(G^x) \xrightarrow{G^x \rightarrow H^x} H_*(H^x) \]

Fact: If \( m_\alpha \) and \( m_\beta \) are isomorphisms and commute with the maps in the persistence modules, then

\[ \text{Pers}(F) = \text{Pers}(G) \]
**Homeomorphism Invariance**

**Thm:** If \( f: Y \rightarrow \mathbb{R} \) is a tame function and \( h: X \rightarrow Y \) is a homeomorphism, then

\[
\text{Pers}(f) = \text{Pers}(f \circ h)
\]

**pf**  
* Exercise  
* Hint: Use Functoriality
Remember These Two Results:

**The Stability Theorem:** \( d_B(\text{Pers}(f), \text{Pers}(g)) \leq \|f - g\|_\infty \)

**Homeomorphism Invariance:** \( \text{Pers}(f) = \text{Pers}(f \circ h) \) for all homeomorphisms \( h \).
New Problem: The Hausdorff distance is not discriminating enough for curves.

Main Issue: We threw out all info about the domain.

One option: \[
\max_{t \in [0,1]} d(f(t), g(t))
\]

... but it's sensitive to the parameterization.
The Fréchet Distance

\[ f, g : [0,1] \rightarrow \mathbb{R}^2 \]

\[
d_F(f, g) := \inf_{h} d_{\text{inf}}(f, g \circ h)
\]

where \( h \) ranges over homeomorphisms \([0,1] \rightarrow [0,1]\)

Easy to generalize to \( f, g : X \rightarrow Y \) where \( X \) is
any compact top. space and \( Y \) is any metric space.

Exercise: Show \( d_F(f, g) \geq d_H(\inf f, \inf g) \)
Let's try persistence on the functions $f, g : X \to Y$

* Need a real-valued function.

Fix $p \in Y$. Let $n_p(x) = d(x, p)$.

$n_p \circ f : X \to \mathbb{R}$  $n_p \circ g : X \to \mathbb{R}$

Idea: Compute

$d_S(f, g) = d_B(\text{Pers}(n_p \circ f), \text{Pers}(n_p \circ g))$
Frechet-Stability

Let $h: X \to X$ be any homeomorphism.

\[ d_h(\text{Pers}(n_o f), \text{Pers}(n_o g)) = d_h(\text{Pers}(n_o f), \text{Pers}(n_o g \circ h)) \]

\[ \leq \| n_o f - n_o g \circ h \|_\infty \]

\[ = \max_{x \in X} |d(p, f(x)) - d(p, g \circ h(x))| \]

\[ \leq \max_{x \in X} d(f(x), g \circ h(x)) \]

\[ = d_\omega(f, g \circ h) \]

So $d_s(f, g) \leq \inf_h d_\omega(f, g \circ h) = d_F(f, g)$
What could go wrong?

Curves are very different, but $d_5$ is zero.

Idea: Use more points.

$$d_{S,p}(f,g) = \max_{p \in P} \, d_B(\text{Pers}(\gamma_p \cdot f), \text{Pers}(\gamma_p \cdot g_0 h))$$
How do we compute $d_5$?

Can’t just compute distances at vertices.
Assume: Euclidean distances.

Piecewise Affine Embedding of a simplicial complex.

The filtration of distances to $p$ on a triangle.

Idea: Use the barycentric subdivision.

Assign each vertex the min distance to $p$ over the corresponding simplex.
Thanks!