Wasserstein Stability for Persistence Diagrams

Etienne Gauthier, Paul Fournier, Dan Timsit

Master Mathématiques Vision Apprentissage

April 1, 2024

Main result: cellular Wasserstein Stability

- K a CW-complex
- $f, g : K \to \mathbb{R}$ two monotonous functions

We have

$$W_p(\mathrm{Dgm}_k(f), \mathrm{Dgm}_k(g))^p \leq \sum_{\dim(\sigma)\in\{k,k+1\}} |f(\sigma) - g(\sigma)|^p \tag{1}$$

$$W_p(\mathrm{Dgm}(f), \mathrm{Dgm}(g)) \leq \|f - g\|_p \tag{2}$$

Applications

- Greyscale images of dim d:

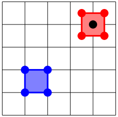

Figure: Two different complexes for images.
Vertices are corners of pixels.
Vertices are pixels.

- Persistence diagram transforms
- Vietoris-Rips complex...

$$W_p(\mathrm{Dgm}(\hat{f}), \mathrm{Dgm}(\hat{g})) \leq \left(\sum_{i=0}^{d} 2^{d-i} \binom{d}{i} \right) \|f - g\|_p$$

Topological summaries

- Positive result for linear function
  - homology rank functions
  - persistence images
  - Betti curves...
- Negative result for persistence landscapes

Define Wasserstein stability at the algebraic level

Definition

A *span* of two persistence modules $\mathcal{A}$ and $\mathcal{B}$ is a triple $(\mathcal{C}, \phi, \psi)$ in a diagram

$$\begin{tikzcd} \mathcal{C} \arrow[r, "\phi"] \arrow[d, "\psi"'] & \mathcal{A} \\ \mathcal{B} & \end{tikzcd}$$

Definition

The *algrebraic p-Wasserstein distance* is defined as $W_p^{\mathrm{alg}}(\mathcal{A}, \mathcal{B}) := \inf \|\ker\phi \oplus \mathrm{coker}\phi \oplus \ker\psi \oplus \mathrm{coker}\psi\|_p$ where the infimum is taken over all possible spans $(\mathcal{C}, \phi, \psi)$ between $\mathcal{A}$ and $\mathcal{B}$.

Recall

$\|\mathcal{A}\|_p = \left(\sum_{x \in \mathrm{Dgm}(\mathcal{A})} l(x)^p\right)^{\frac{1}{p}}$ and $\mathrm{coker}(\phi) = \mathcal{A}/\mathrm{Im}(\phi)$

Lemmas (without their dual formulations)

Lemma

$\mathcal{A} \hookrightarrow \mathcal{B} \Rightarrow \|\mathcal{A}\|_p \leq \|\mathcal{B}\|_p$

Definition

For $t \in [0,1]$, we define the *death-birth interpolation* of $\mathcal{A}$ by:
$\mathcal{A}_t := \bigoplus_{x \in \mathrm{Dgm}(\mathcal{A})} \mathbb{I}\left\{b(x), tb(x) + (1-t)d(x)\right\}$

Lemma

Given $\phi : \mathcal{A} \hookrightarrow \mathcal{B}$, $\forall t \in [0,1]$, $\exists \mathcal{B}_t$ such that:

$$\begin{tikzcd}
\mathcal{A} \arrow[r, hook] \arrow[d, two heads] & \mathcal{B} \arrow[d, two heads] \\
\mathcal{A}_t \arrow[r, hook] \arrow[d, two heads] & \mathcal{B}_t \arrow[d, two heads] \\
0 \arrow[r, hook] & \operatorname{coker} \phi
\end{tikzcd}$$

commutes. And an exact sequence

$$0 \longrightarrow \mathcal{A} \hookrightarrow \mathcal{B} \twoheadrightarrow \mathcal{C} \longrightarrow 0$$

yields another exact sequence

$$0 \longrightarrow \mathcal{A}_t \hookrightarrow \mathcal{B}_t \twoheadrightarrow \mathcal{C} \longrightarrow 0$$

Lemma

Given a monomorphism $f : \mathcal{A} \hookrightarrow \mathcal{B}$ between pfd modules where coker f is finitely generated , we have: $W_p(\mathrm{Dgm}(\mathcal{A}), \mathrm{Dgm}(\mathcal{B})) \leq \|\mathrm{coker}\ f\|_p$

Theorem

Given

$$0 \longrightarrow \mathcal{A} \hookrightarrow \mathcal{B} \twoheadrightarrow \mathcal{C} \longrightarrow 0$$

then $W_p(\mathrm{Dgm}(\mathcal{A}), \mathrm{Dgm}(\mathcal{B})) \leq \|\mathcal{C}\|_p$ and $W_p(\mathrm{Dgm}(\mathcal{B}), \mathrm{Dgm}(\mathcal{C})) \leq \|\mathcal{A}\|_p$.

Corollary

Given $f : \mathcal{A} \to \mathcal{B}$, $W_p(\mathrm{Dgm}(\mathcal{A}), \mathrm{Dgm}(\mathcal{B})) \leq \|\ker f \oplus \mathrm{coker}\ f\|_p$

$W_p^{alg} = W_p$

Theorem

Given $\mathcal{A}$, $\mathcal{B}$ and a span $(\mathcal{C}, \phi, \psi)$ we have for $p < \infty$:
$W_p(\mathrm{Dgm}(\mathcal{A}), \mathrm{Dgm}(\mathcal{B}))^p \leq \|\mathrm{ker}\ \phi\|_p^p + \|\mathrm{coker}\ \phi\|_p^p + \|\mathrm{ker}\ \psi\|_p^p + \|\mathrm{coker}\ \psi\|_p^p$.

Theorem

Let $\mathcal{M}$ be a matching between two diagrams corresponding to $\mathcal{A}$ and $\mathcal{B}$.
For $p < \infty$, $\mathcal{M}$ induces a span $(\mathcal{C}, \phi, \psi)$ such that:
$\sum_{(x,y)\in\mathcal{M}} |b(x) - b(y)\|^p + |d(x) - d(y)\|^p =$
$\|\mathrm{ker}\ \phi\|_p^p + \|\mathrm{coker}\ \phi\|_p^p + \|\mathrm{ker}\ \psi\|_p^p + \|\mathrm{coker}\ \psi\|_p^p$.

Conclusion: the algebraic and diagram Wasserstein distances agree.

$\|\mathcal{A} \oplus \mathcal{C}\|_p \leq \|\mathcal{B}\|_p$

Theorem

Given a short exact sequence of persistence diagrams:

$$0 \longrightarrow \mathcal{A} \hookrightarrow \mathcal{B} \twoheadrightarrow \mathcal{C} \longrightarrow 0$$

then $\|\mathcal{A} \oplus \mathcal{C}\|_p \leq \|\mathcal{B}\|_p$.

Corollary

If $f, g : K \to \mathbb{R}$, and if $\mathcal{F}, \mathcal{G}$ are the associated persistence modules and are pointwise finite dimensional persistence modules, then:

$$W_p(\mathrm{Dgm}(f), \mathrm{Dgm}(g)) \leq \|f - g\|_p \tag{3}$$

Properties of the distance function

Theorem

The p-Wasserstein algebraic norm function is a pseudo-norm on the set of pointwise finite dimensional persistence modules with the following properties:

- $W_p(0) = 0$
- $\mathcal{A} \hookrightarrow \mathcal{B} \Rightarrow \|\mathcal{A}\|_p \leq \|\mathcal{B}\|_p$
- Subadditivity
- Triangle inequality thanks to a lemma using spans and the algebraic definition of the norm