Persistent homology and statistical inference: Persistence Landscapes

Peter Bubenik

Cleveland State University

January 5, 2012

Joint Mathematical Meetings 2012
Boston, MA
An applied topology pipeline

Raw data $\xrightarrow{\text{Encode and preprocess}}$ Cleaned data

Geometry $\xrightarrow{\text{Geometric object}}$ Algebraic topology $\xrightarrow{\text{Persistent homology}}$ Descriptor

Algebraic diagram
The usual descriptor
Suppose we have calculated a sequence of descriptors:

**Question**

What is the mean? What is the standard deviation? Can we use it for hypothesis testing?
Suppose we have calculated a sequence of descriptors:

**Question**

What is the mean? What is the standard deviation? Can we use it for hypothesis testing?

**A solution**

We define a new descriptor: *the persistence landscape*. It is a function from $\mathbb{R}^2$ to $\mathbb{R}$. 
Persistence Landscape example
Let $M$ be a vector space diagram indexed by $\mathbb{R}$.
For all $a \leq b$, we have a linear transformation $M(a) \to M(b)$.
Define $\beta^{a,b}(M) = \dim(\text{im}(M(a) \to M(b)))$.

**Definition**

$$\lambda(M)(x, t) = \begin{cases} 
\sup(s \mid \beta^{t-s, t+s}(M) \geq x) & \text{if } 0 < x \leq \dim M(t) \\
0 & \text{otherwise.}
\end{cases}$$
Motivation
Persistence Landscape
Statistics
Examples

## Persistence Landscape

Let $M$ be a vector space diagram indexed by $\mathbb{R}$. For all $a \leq b$, we have a linear transformation $M(a) \to M(b)$. Define $eta_{a,b}(M) = \dim(\text{im}(M(a) \to M(b)))$.

### Definition

\[
\lambda(M)(x, t) = \begin{cases} 
\sup(s \mid \beta_{t-s, t+s}(M) \geq x) & \text{if } 0 < x \leq \dim M(t) \\
0 & \text{otherwise.}
\end{cases}
\]

### Theorem

Let $\lambda_B$ denote the persistence landscape corresponding to a barcode $B$. Then,

\[
\|\lambda_B\|_2^2 = \frac{1}{12} \text{pers}_3(B).
\]
The space of persistence landscapes

**Definition**

Let PL be the set of functions that “look like persistence landscapes” (a “decreasing stack of 1-Lipschitz functions”). Let $PL^2 = PL \cap L^2$ together with the metric from $L^2$. 

**Theorem**

$PL^2$ is complete and separable, i.e. it is a Polish space.

**Peter Bubenik**

**Persistence Landscapes**
The space of persistence landscapes

Definition
Let PL be the set of functions that “look like persistence landscapes” (a “decreasing stack of 1-Lipschitz functions”). Let $PL^2 = PL \cap L^2$ together with the metric from $L^2$.

Theorem
$PL^2$ is complete and separable, i.e. it is a Polish space.
**Fréchet mean, variance**

**Definition**

Given $x_1, \ldots, x_n$ in a metric space. The Fréchet mean is the point $y$ that minimizes

$$\sum_{i=1}^{n} d(y, x_i)^2.$$ 

The corresponding sum is called the Fréchet variance.

Replacing summation with integration we define the Fréchet mean and variance of a probability measure.
**Fréchet mean, variance**

**Definition**

Given $x_1, \ldots, x_n$ in a metric space. The Fréchet mean is the point $y$ that minimizes

$$\sum_{i=1}^{n} d(y, x_i)^2.$$ 

The corresponding sum is called the Fréchet variance.

Replacing summation with integration we define the Fréchet mean and variance of a probability measure.

**Theorem**

In $\text{PL}^2$, the Fréchet mean is given by the pointwise mean, and the Fréchet variance is the integral of the pointwise variances.
Let $\lambda_1, \ldots, \lambda_n$ be a sample of persistence landscapes drawn from a probability measure with Fréchet mean $h$. Let $\bar{\lambda}_n$ be the pointwise mean of the sample.
Estimator for the Fréchet mean

Let $\lambda_1, \ldots, \lambda_n$ be a sample of persistence landscapes drawn from a probability measure with Fréchet mean $h$. Let $\bar{\lambda}_n$ be the pointwise mean of the sample.

For all $x$, $t$,

**Strong Law of Large Numbers**

$$\bar{\lambda}_n(x, t) \xrightarrow{a.s.} h(x, t).$$

**Central Limit Theorem**

$$\sqrt{n} \left( \bar{\lambda}_n(x, t) - h(x, t) \right) / \sigma \xrightarrow{d} N(0, 1).$$
Torus: point cloud
Torus: filtered simplicial complex

see rendering from plex_viewer
Torus: persistence landscape in dimension 1
Torus: persistence landscape in dimension 1

Now repeat 10 times.

Average pointwise.
Torus: average persistence landscape in dimension 1
Motivation
Persistence Landscape
Statistics
Examples

Torus: average persistence landscape in dimension 1

±2 standard deviations

[Diagrams showing the persistence landscape]
Torus vs Sphere: point clouds
Torus vs Sphere: average persistence landscapes dim 0
Torus vs Sphere: average persistence landscapes dim 1
Torus vs Sphere: average persistence landscapes dim 2
Question

Is there a statistically significant difference?

Answer

Using a two sample t test at the 0.01 significance level:

<table>
<thead>
<tr>
<th></th>
<th>dim 0</th>
<th>dim 1</th>
<th>dim 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>peak 1</td>
<td>no</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>peak 2</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
</tbody>
</table>
Noisy Torus vs Sphere: point clouds
Noisy Torus vs Sphere: avg persistence landscapes dim 0
Noisy Torus vs Sphere: avg persistence landscapes dim 1
Noisy Torus vs Sphere: avg persistence landscapes dim 2
Noisy Torus vs Noisy Sphere

Question

Is there a statistically significant difference?

Answer

Using a two sample t test at the 0.01 significance level:

<table>
<thead>
<tr>
<th></th>
<th>dim 0</th>
<th>dim 1</th>
<th>dim 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>peak 1</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>peak 2</td>
<td>no</td>
<td>yes</td>
<td>no</td>
</tr>
</tbody>
</table>

Peter Bubenik
Persistence Landscapes
Future directions

- Stability: \( \|\lambda(M) - \lambda(N)\|_2 \leq K \, d(M, N) \) ?
- Fréchet median
- Uniform convergence, bootstrap
- Asymptotic persistence landscapes
- Applications!
## Torus vs Sphere: p values

<table>
<thead>
<tr>
<th>peak</th>
<th>dim 0</th>
<th>dim 1</th>
<th>dim 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>peak 1</td>
<td>0.8807</td>
<td>0.0000</td>
<td>0.7357</td>
</tr>
<tr>
<td>peak 2</td>
<td>0.7547</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>peak 3</td>
<td>0.4038</td>
<td>0.0002</td>
<td>0.0000</td>
</tr>
<tr>
<td>peak 4</td>
<td>0.7163</td>
<td>0.0032</td>
<td>0.0000</td>
</tr>
<tr>
<td>peak 5</td>
<td>0.2837</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>
## Noisy Torus vs Noisy Sphere: p values

<table>
<thead>
<tr>
<th></th>
<th>dim 0</th>
<th>dim 1</th>
<th>dim 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>peak 1</td>
<td>0.3246</td>
<td>0.0077</td>
<td>0.0000</td>
</tr>
<tr>
<td>peak 2</td>
<td>0.4967</td>
<td>0.0032</td>
<td>0.3217</td>
</tr>
<tr>
<td>peak 3</td>
<td>0.8132</td>
<td>0.0000</td>
<td>0.3306</td>
</tr>
</tbody>
</table>