Topological Data Analysis

21 November 2019

5 Persistence diagrams and landscapes

5.1 Persistence diagrams

Suppose given a persistence module (V, π) over a field $\mathbb{F}$ with normal form

$$V \cong \bigoplus_{i=1}^{N} \mathbb{F}(I_i)^{m_i} \tag{5.1}$$

where $I_i = [b_i, d_i)$ or $I_i = [b_i, \infty)$ for each i. We assume that $I_i \neq I_j$ if $i \neq j$ and $m_i \geq 1$ for all i.

The *persistence diagram* for (V, π) has a point (b_i, d_i) in a coordinate plane for each bounded interval $[b_i, d_i)$ in its normal form (5.1). Thus a point (b_i, d_i) in a persistence diagram denotes a basis vector of V_* with birth parameter b_i and death parameter d_i, where $V_* = V_{a_0} \oplus \cdots \oplus V_{a_k}$ if $\{a_0, \ldots, a_k\}$ is the spectrum of (V, π).

The rays $[b_i, \infty)$ are represented as points (b_i, y_∞) where $y_\infty = a_k$ is the largest value in the spectrum of (V, π). The multiplicites m_i are usually depicted by increasing the size of the corresponding dots in the picture. It is also customary to include the diagonal $b = d$ in persistence diagrams, and view its points as having infinite multiplicity.

5.2 Landscapes

The *landscape* of a persistence module (V, π) over a field $\mathbb{F}$ is a collection of piecewise linear continuous functions $\lambda_k \colon \mathbb{R} \to \mathbb{R}$ defined as follows. For $k \geq 1$, let

$$\lambda_k(t) = \sup\{x \geq 0 \mid \beta_{t-x,\,t+x} \geq k\}$$

where we denote

$$\beta_{t-x,\,t+x} = \dim_{\mathbb{F}} \operatorname{Im}(\pi_{t-x,\,t+x}),$$

and define $\lambda_k(t) = 0$ if there is no x such that $\beta_{t-x,\,t+x} \geq k$.

We next give a more useful description of landscape functions. For real numbers $b < d$, consider the function

$$\Lambda_{(b,d)}(t) = \max\{0, \min\{t - b, d - t\}\}.$$

If $\{(b_i, d_i)\}_{i=1}^{N}$ is the set of points (with their respective multiplicities m_i) in the corresponding persistence diagram, we have that

$$\lambda_k(t) = \text{kmax}\{\Lambda_{(b_i,d_i)}(t)\}_{i=1}^{N}$$

where kmax returns the kth largest value in a given set of real numbers, counted with their multiplicities, or 0 if there is no kth largest value.

For simplicity, in landscape pictures arising from point clouds, the points (b_i, y_∞) in the persistence diagram corresponding to infinite rays may be omitted.

Short exercise

(1) Draw a persistence diagram and a landscape for the following persistence module over a field $\mathbb{F}$:

$$\mathbb{F}[0,\infty) \oplus \mathbb{F}[0,2) \oplus \mathbb{F}[1,4) \oplus \mathbb{F}[1,2) \oplus \mathbb{F}[3,5).$$

Longer exercise

(1) Use your favorite software to draw a persistence diagram and landscape functions for the Vietoris–Rips persistence module of the following point cloud. The R package TDA is recommended [B. T. Fasy, J. Kim, F. Lecci, C. Maria, Introduction to the R package TDA, 2015, HAL Id: hal-01113028 (2015), https://hal.inria.fr/hal-01113028].

$$\begin{array}{cccc}
(0.01, 0.93) & (0.05, 1.12) & (0.28, 0.29) & (0.29, 1.77) \\
(0.32, 0.27) & (0.33, 1.64) & (0.90, 1.95) & (0.92, 0.15) \\
(0.94, 1.98) & (1.21, 0.33) & (1.65, 1.71) & (1.67, 0.31) \\
(1.81, 1.72) & (1.88, 0.91) & (1.97, 0.98) & (1.99, 0.95)
\end{array}$$