



R Package TDA for Statistical Inference on Topological Data Analysis

Jisu KIM (jisuk1@andrew.cmu.edu)

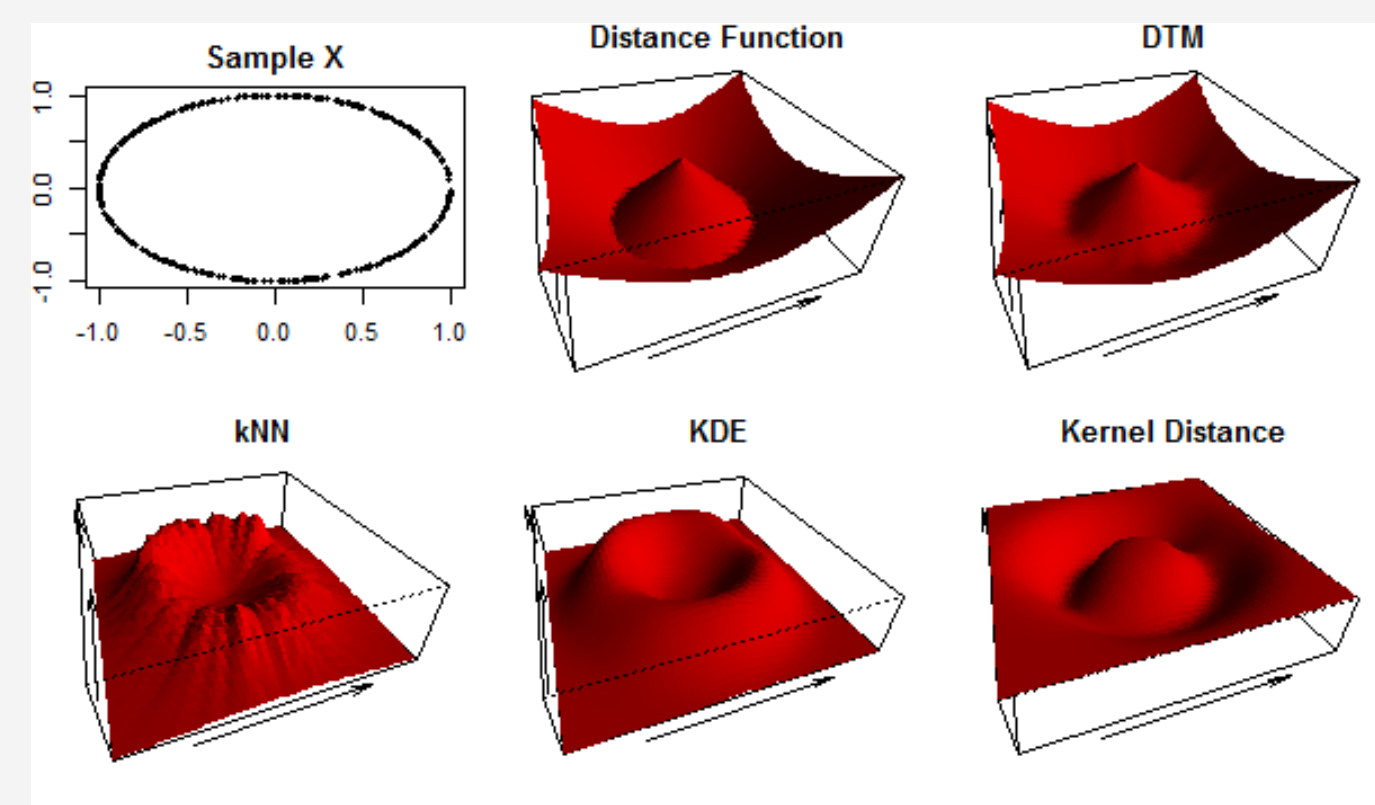
Carnegie Mellon University

R Package TDA

- website: <https://cran.r-project.org/web/packages/TDA/index.html>
- Author: Brittany Terese Fasy, Jisu Kim, Fabrizio Lecci, Clément Maria, and Vincent Rouvreau
- The salient topological features of data can be quantified with persistent homology.
- The R package TDA provides an R interface for C++ library GUDHI / Dionysus / PHAT, including functions for computing the persistent homology.
- The R package TDA also computes the confidence band to distinguish significant topological features from noisy features in the persistent homology.
- The R package TDA also includes an algorithm for the cluster tree that corresponds to the density clustering.

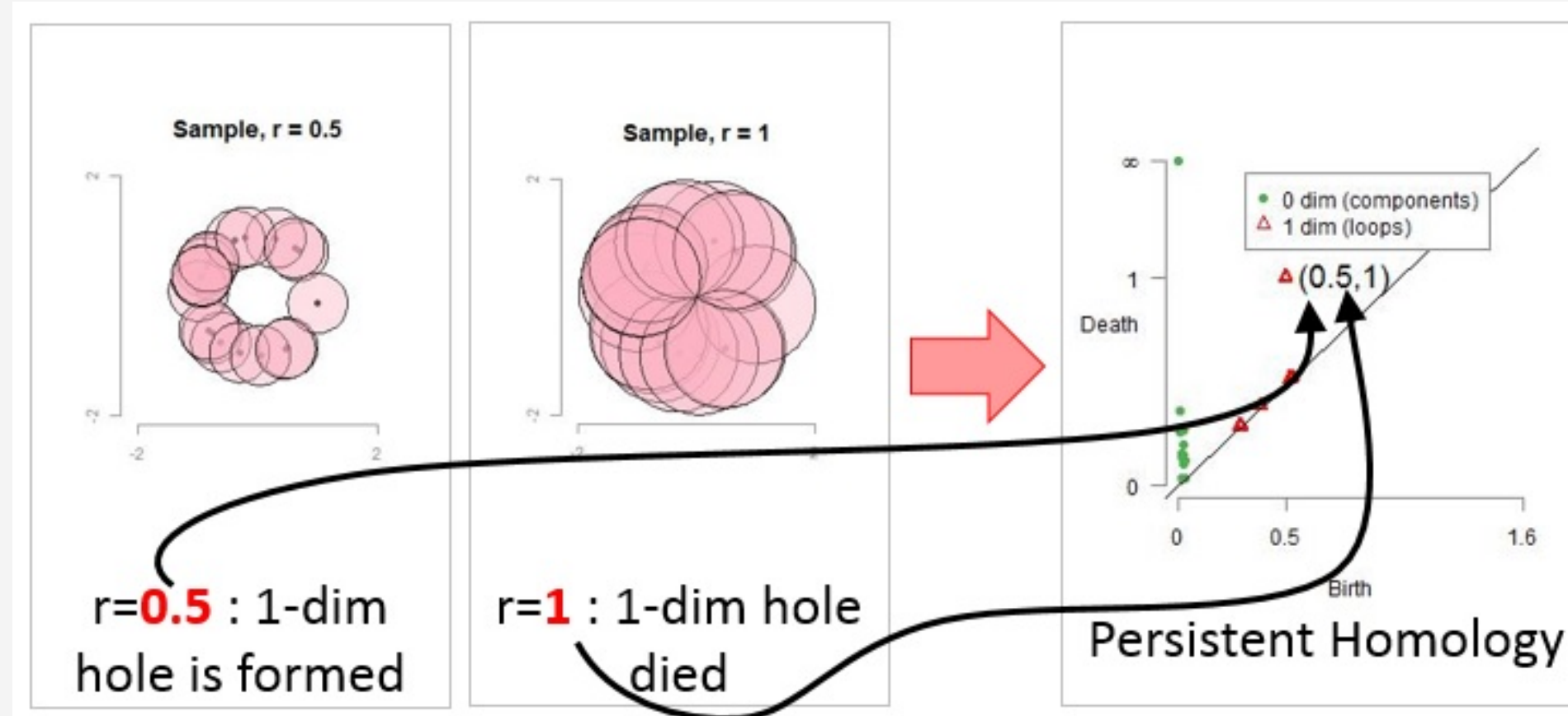
Distance Functions and Density Estimators

- The R package TDA provides various distance functions and density estimators computed over a grid of points.

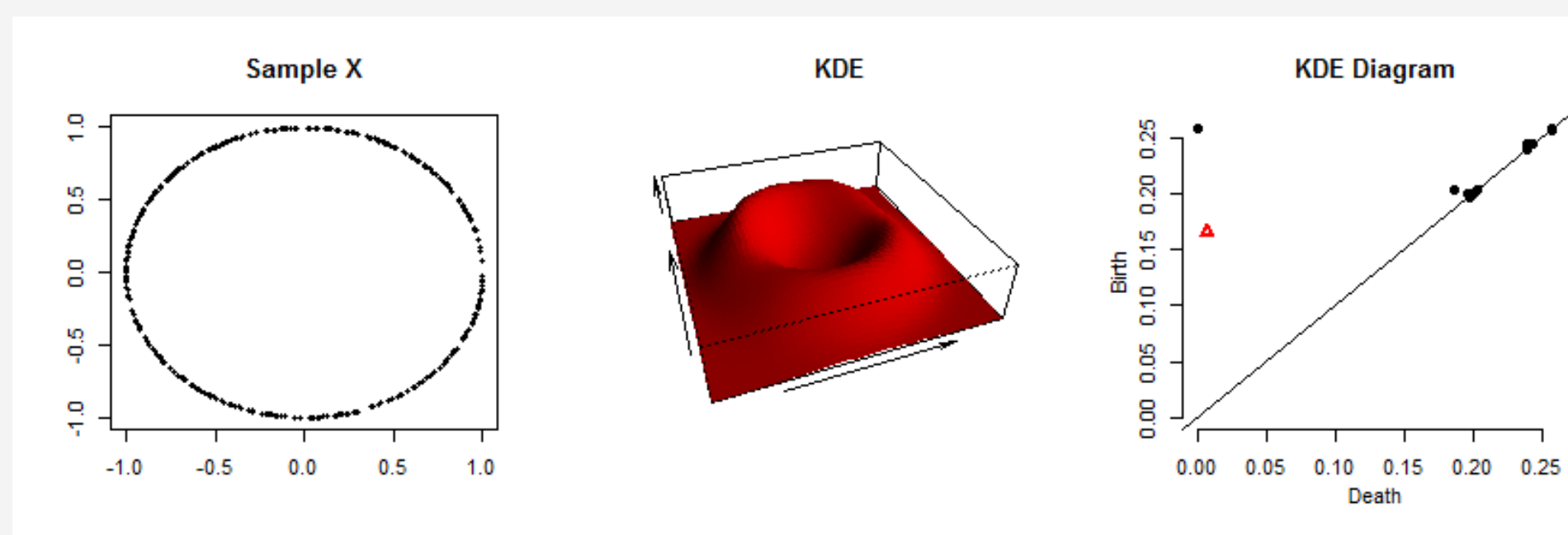


Persistent Homology

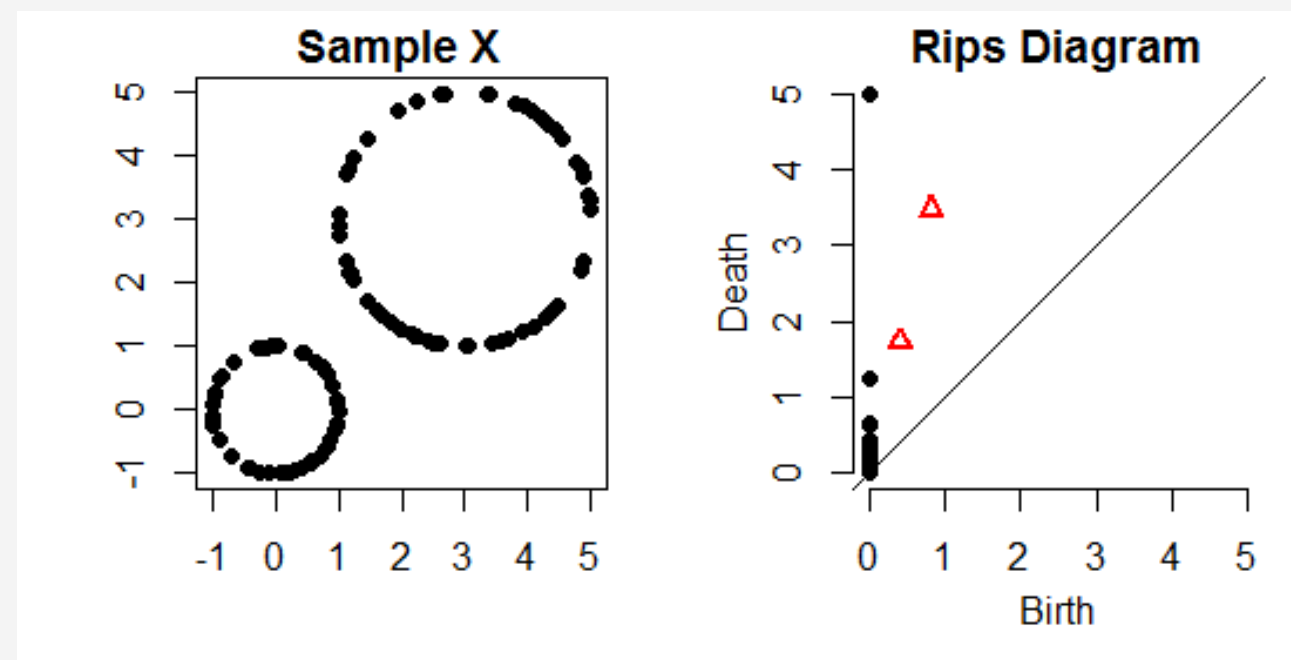
- Persistence homology computes homologies on collection of sets, and tracks when topological features are born and when they die.



- The function gridDiag() computes the persistent homology of sublevel (and superlevel) sets of the input function.

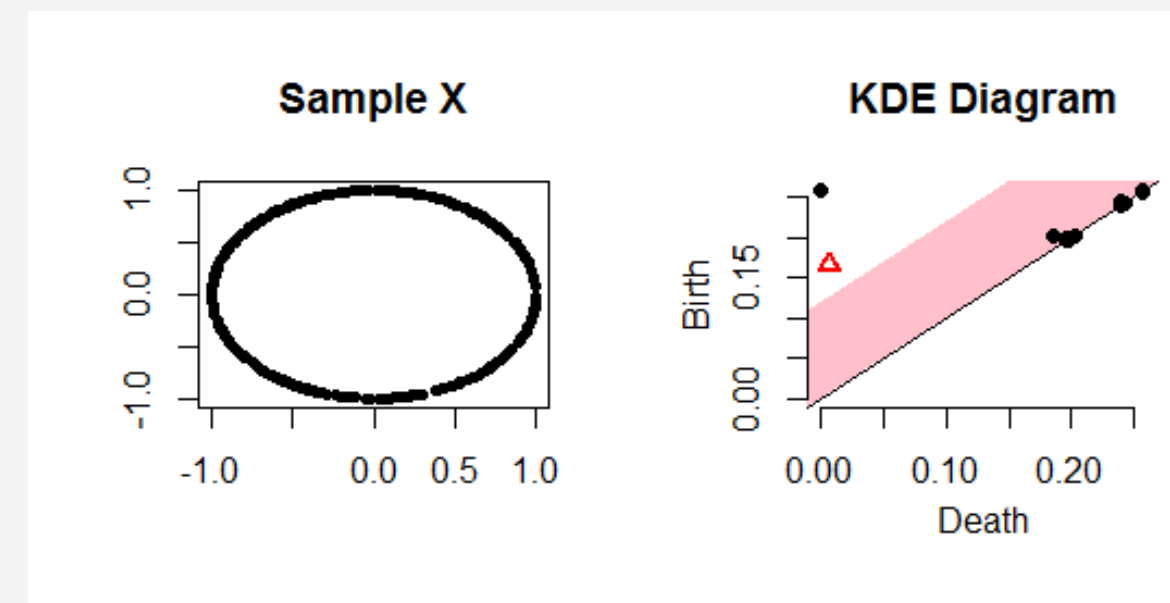


- The function ripsDiag() computes the persistent homology of the Rips filtration built on top of a point cloud.



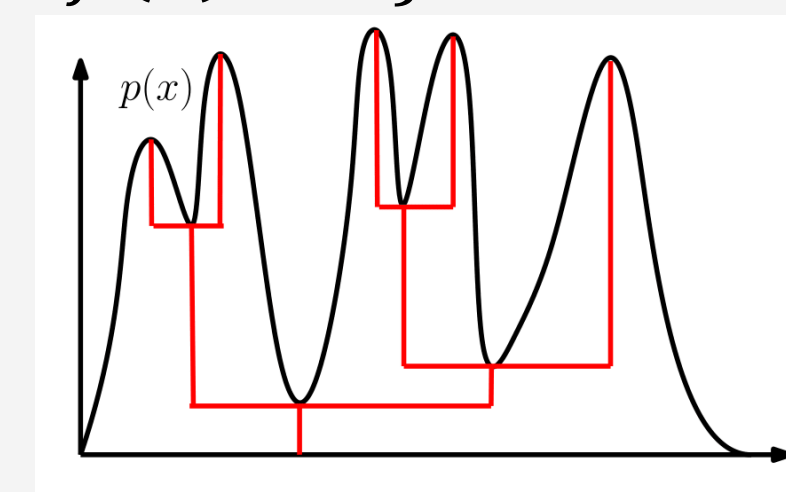
Statistical Inference on Persistent Homology

- The function bootstrapBand() computes $(1-\alpha)$ bootstrap confidence band.



Cluster Tree

- For any function f , the cluster tree of f is a function T_f , where $T_f(\lambda)$ is the set of the connected components of the upper-level set $\{x: f(x) \geq \lambda\}$.



- The function clusterTree() computes the cluster tree.

