R Package TDA:
Statistical Tools for Topological Data Analysis

Jisu Kim, Carnegie Mellon University

Abstract

This poster gives an introduction to the R package TDA, which provides some tools for Topological Data Analysis. The salient topological features of data can be quantified with persistent homology. R package TDA provide an R interface for the efficient algorithms of the C++ libraries GUDHI, Dionysus, and PHAT, including functions for computing 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

- website: https://cran.r-project.org/web/packages/TDA/index.html
- Author: Brittany Terese Fasy, Jisu Kim, Fabrizio Lecci, Clément Maria, and Vincent Rouvreau
- R is a programming language for statistical computing and graphics.
- R has short development time, while C/C++ has short execution time.
- R package TDA provides an R interface for C++ library GUDHI / Dionysus / PHAT.
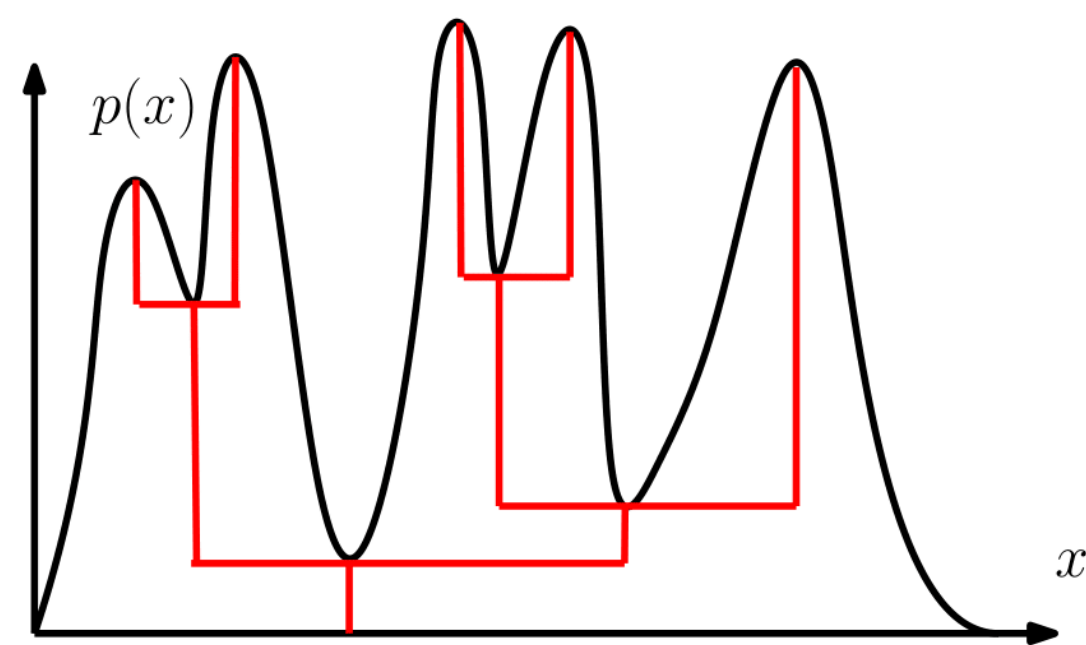
Distance Functions and Density Estimators

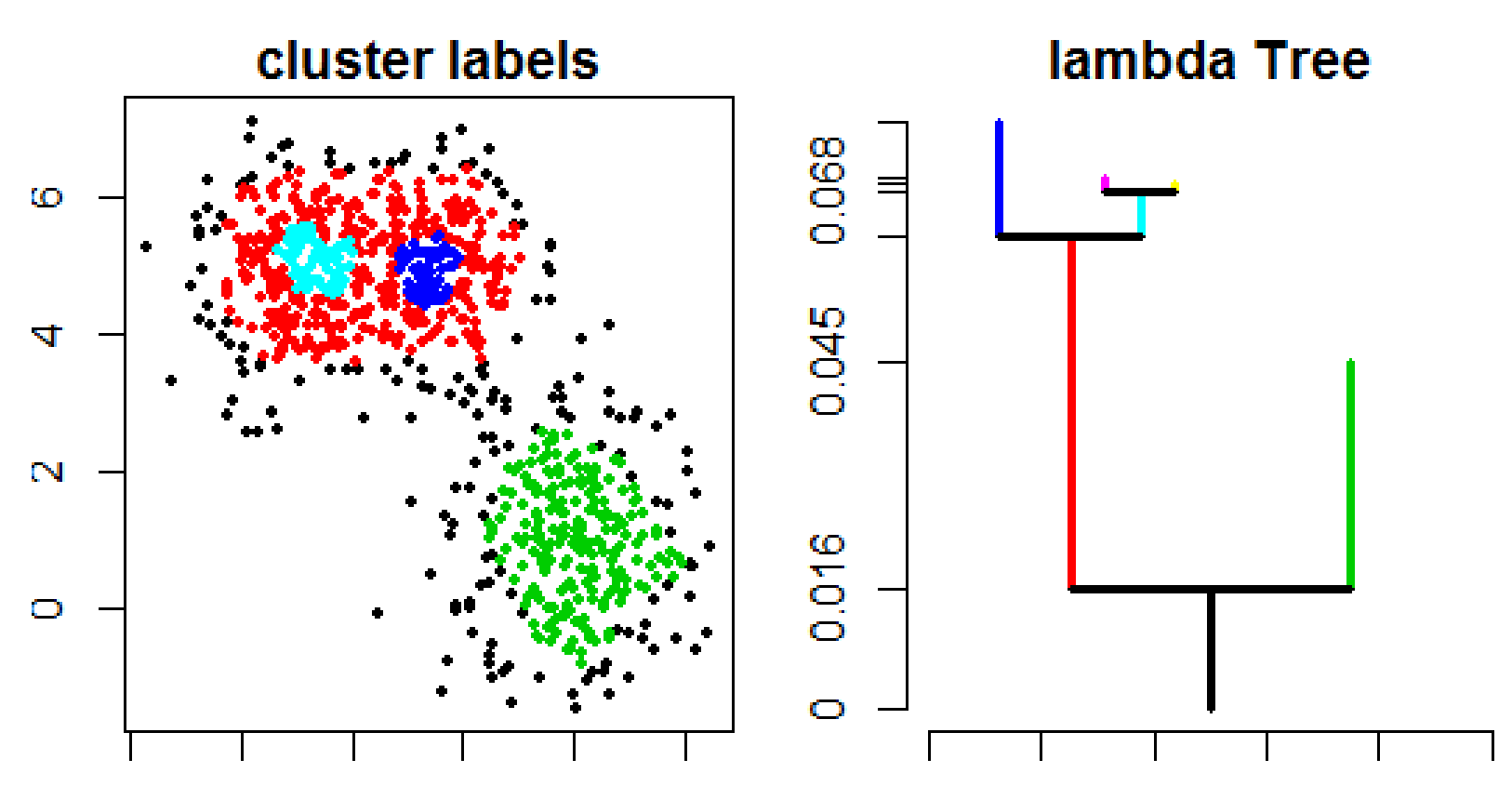- R package TDA provides various distance functions and density estimators computed over a grid of points.
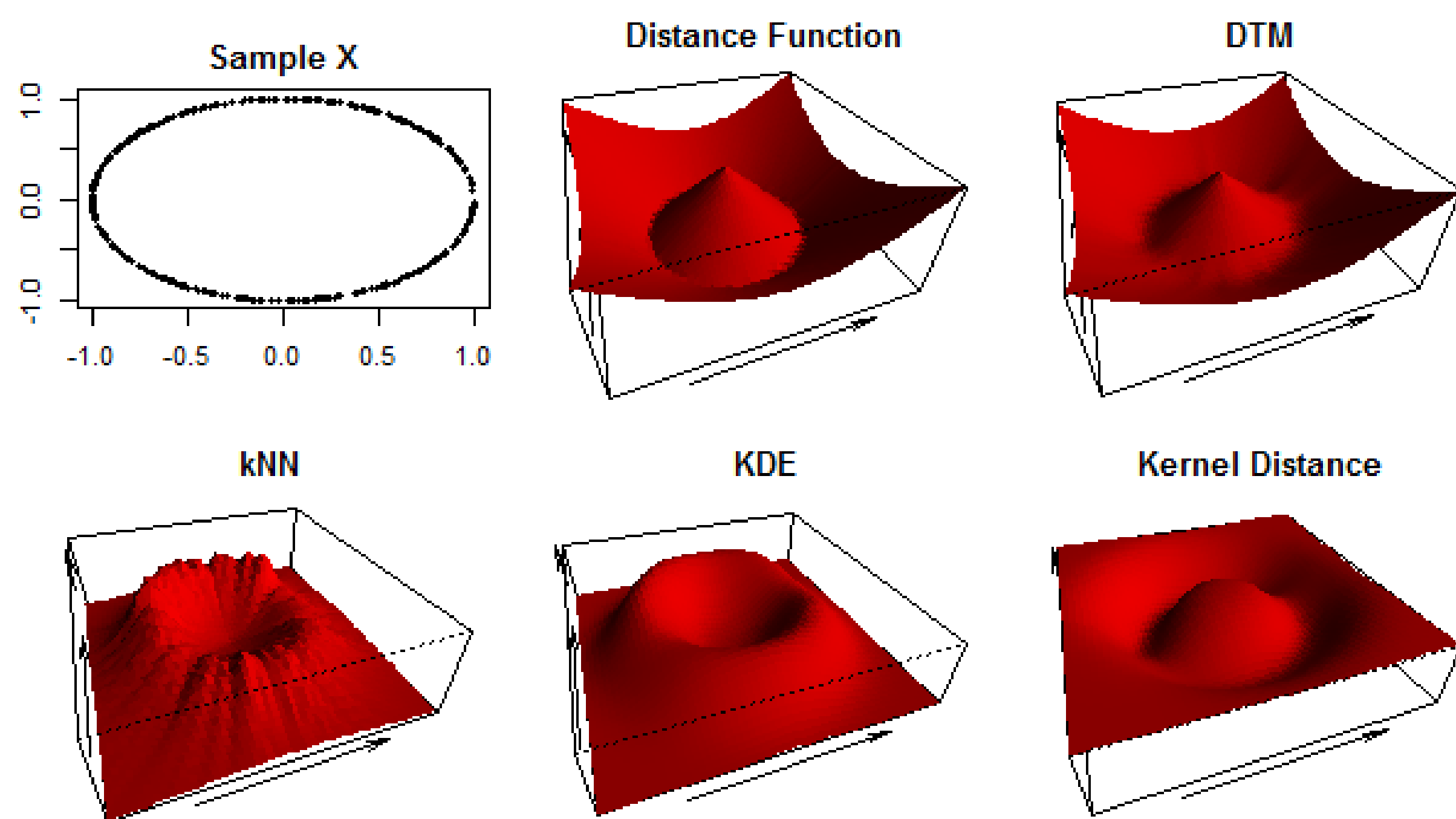
Cluster Trees

- 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 $\{x: f(x) \geq \lambda\}$.
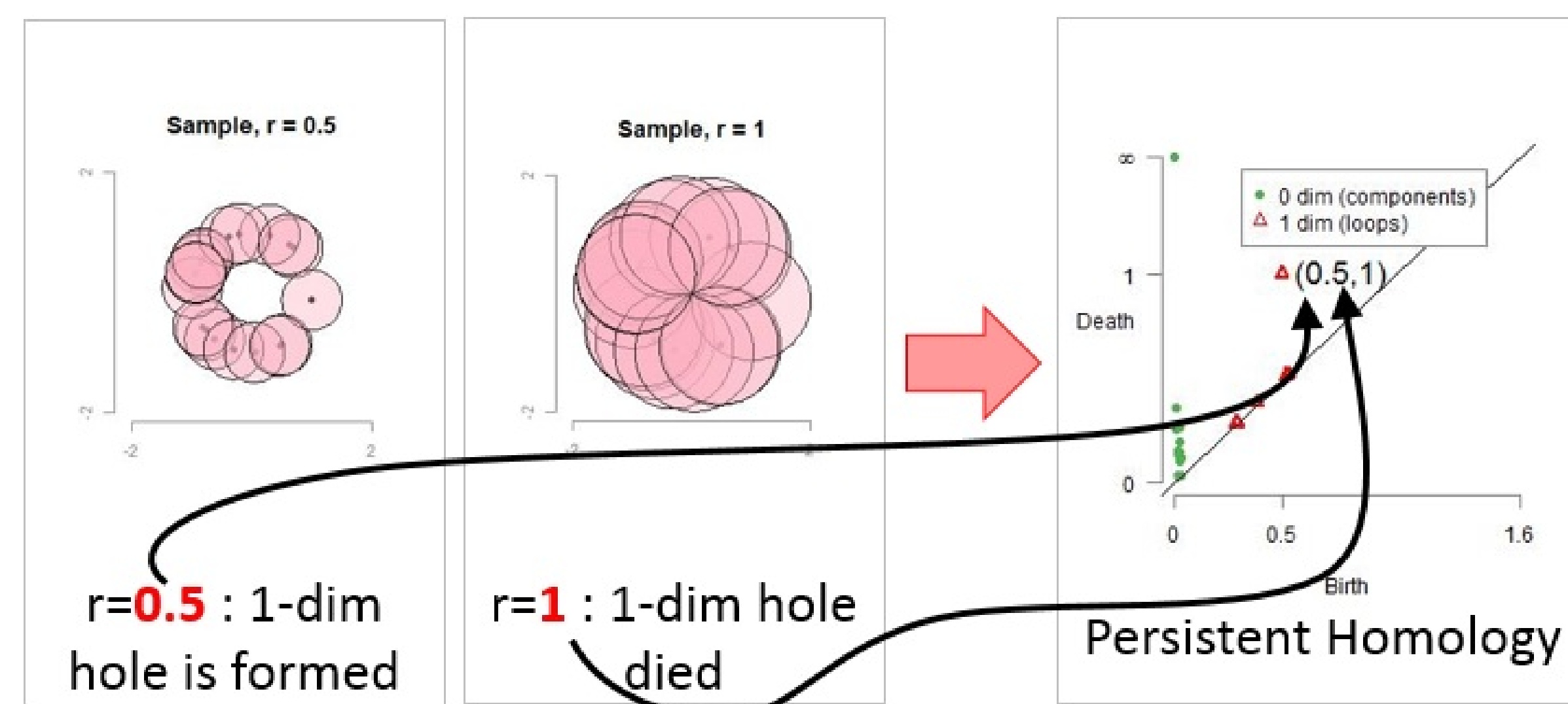

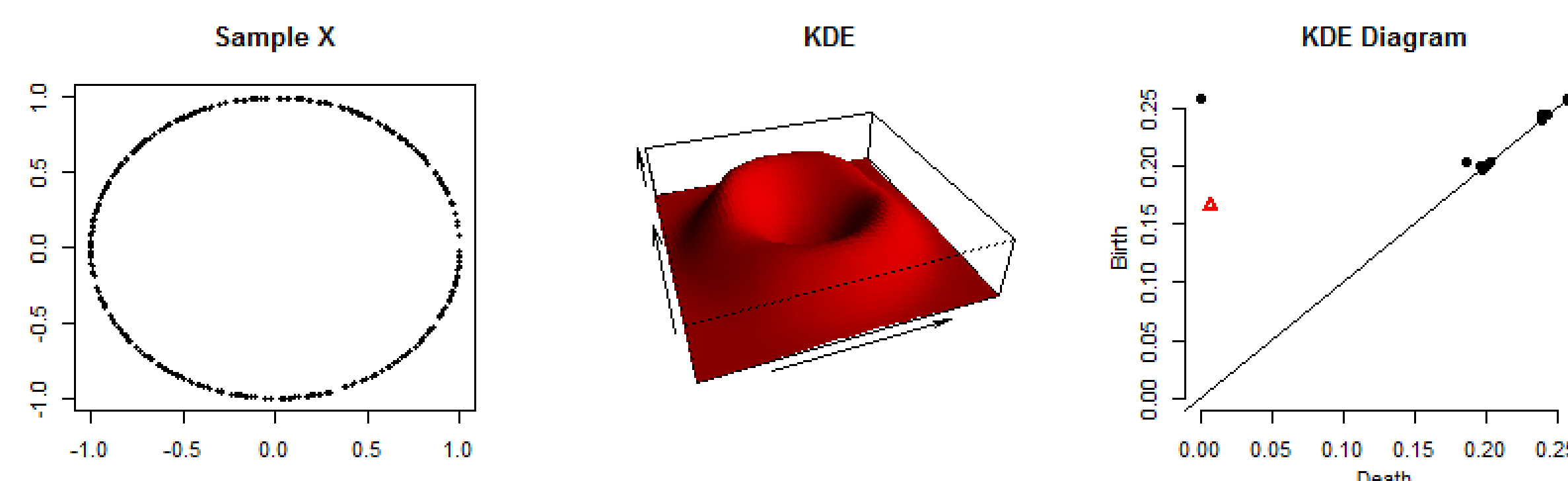- The function clusterTree() computes the cluster tree.


Persistent Homology

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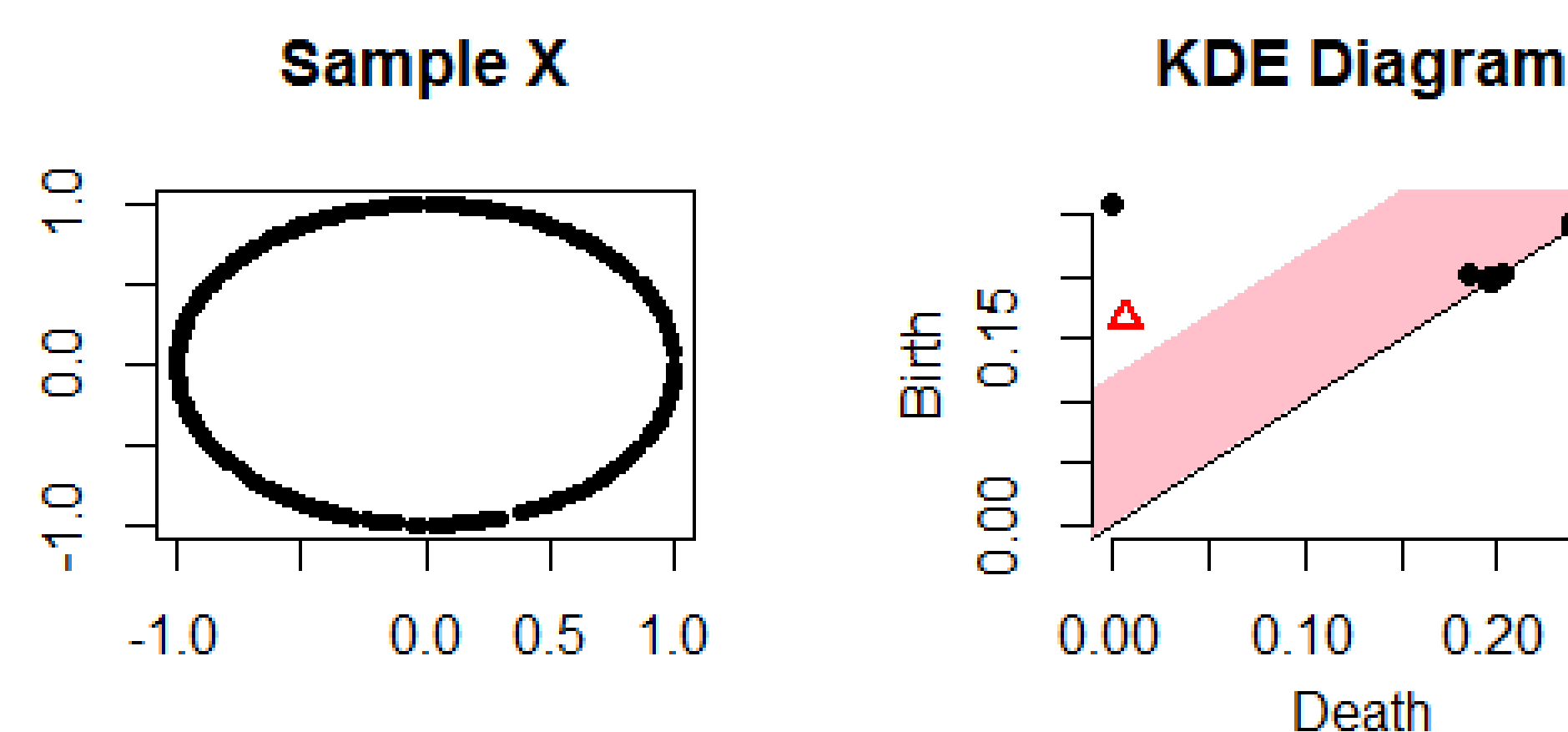

- The function gridDiag() computes the persistent homology of sublevel (and superlevel) sets of the input function.

- The function ripsDiag() computes the persistent homology of the Rips filtration built on top of a point cloud.
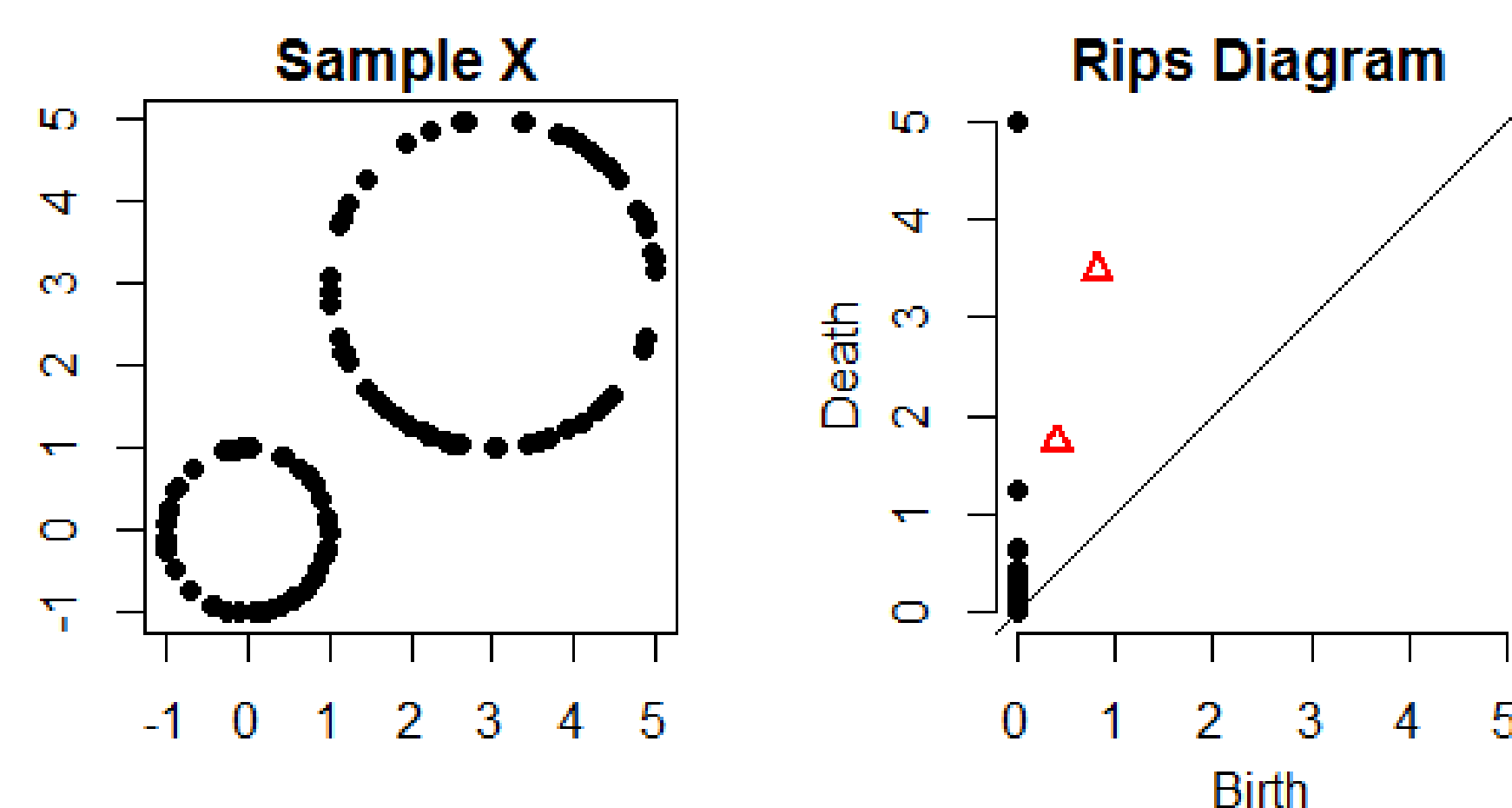
- The function bootstrapBand() computes (1-α) bootstrap confidence band.

Reference

Brittany Terese Fasy, Jisu Kim, Fabrizio Lecci, Clément Maria, and Vincent Rouvreau, Introduction to R Package TDA