Introduction to Tutorial on the R package TDA

Jisu Kim

Brittany T. Fasy, Jisu Kim, Fabrizio Lecci, Clément Maria, Vincent Rouvreau

R is a programming language for statistical computing and graphics.

- Many packages for statistical computing.
- Easy to make (interactive) plots.
- Script language.
- Easy to use.
- Functional language.

- ... but slow.

R has short development time, while C/C++ has short execution time.

R	C / C++
• Many packages	• Many packages
• Easy to make plots	• Difficult to make plots
• Script language	• Compiler language
• Easy to use	• Difficult to use
• Functional language	• Not functional
• Slow	• Fast

R
short development time

C / C++
short execution time

GUDHI, Dionysus, and PHAT are C++ library for Topological Data Analysis.

R
short development time

GUDHI / Dionysus / PHAT
short execution time

R package TDA bridges between R and C++ library GUDHI/Dionysus/PHAT.

R
short development time

TDA

GUDHI / Dionysus / PHAT
short execution time

R package TDA provides an R interface for GUDHI, Dionysus, and PHAT.

R
short development time

TDA

GUDHI / Dionysus / PHAT
short execution time

R package TDA provides tools for Topological Data Analysis.

You can compute:

- common distance functions and density estimators
- persistent homology of the Rips filtration
- persistent homology of sublevel sets of a function over a grid
- confidence band for the persistence diagram
- cluster density trees for density clustering