Applications of Topological Data Analysis in Economics

Christopher Shultz

September 19, 2024

**Abstract**

Topological Data Analysis (TDA) is an emerging tool that offers unique insights for analyzing complex, high-dimensional data, which are becoming increasingly common in economic research. This paper introduces TDA to the economics literature, providing an accessible overview of its core concepts and methods, with a focus on how it differs from traditional econometrics. We also present a critical review of the existing applications of TDA in economics, focusing on areas where it has demonstrated value, such as financial markets, time series dynamics, and systemic risk quantification. Additionally, we evaluate the limitations of TDA, particularly regarding its computational complexity, interpretability, and integration with standard techniques. By offering a nuanced view of TDA’s strengths and weaknesses, this paper aims to serve as a resource for economists seeking to apply TDA to their own research.

# 1: Introduction

As economic data grows increasingly complex, traditional methods may struggle to capture the full extent of embedded relationships, particularly when non-linearity and high-dimensionality are present. Topological Data Analysis (TDA) offers an alternative by focusing on the shape and structure of data, providing a new lens that may be useful for prediction and inference. Unlike conventional approaches that may impose rigid assumptions or parsimony, TDA provides flexible, nonparametric tools for extracting information about datasets that may otherwise remain hidden.

TDA has gained traction in fields such as biology and neuroscience, showing success in uncovering topological features like clusters, holes, and voids in data, offering novel insights. Despite these successes, TDA remains a relatively fringe topic in the economics literature. This paper aims to bridge that gap by introducing TDA to the economics community and exploring its potential as a complementary tool for economic research. We discuss how TDA can add value to the analysis of high-dimensional data, and discuss examples of where it has shown promise such as the understanding of complex financial networks, examination of economic cycles, and commodities markets.

TDA provides a robust framework for understanding the geometric structure of data, particularly useful when examining datasets with complex multi-dimensional relationships, nonlinearity, and other violations of conventional modeling assumptions. The broader goal of this paper is to raise awareness among economists about the potential applications of TDA and provide a roadmap for integrating TDA into economics research.

Grounded in algebraic topology, TDA provides a mechanism to capture and analyze the underlying “shape” of data, first by representing data as a “point cloud” and then constructing a “simplicial complex”, which is an approximation of the underlying topological structure. TDA then extracts features from this structure, such as connected components, loops, and voids, which can provide valuable information about the underlying structure and dynamics of the dataset.

The core tool of TDA is *persistent homology*, a method for identifying and tracking the emergence and disappearance of these topological features across different scales or resolutions. This allows researchers to focus on the most significant features in the data, filtering out noise and capturing only essential structures. A variety of methods exist to visualize the persistence of topological features therein, such as persistence diagrams and the mapper representation.

Compared to traditional econometric approaches, TDA avoids the imposition of assumptions about the underlying structure of the data, allowing for the identification of nonlinearities, hidden patterns, and high-dimensional features. Overall, we view TDA as a *complementary* approach to traditional methods which can be especially useful in the case of complex or high-dimensional data. Despite this promise, TDA is not without challenges, such as computational complexity and the interpretation of topological features.

The aim of this paper is to provide an accessible but comprehensive introduction to TDA for the economics community. We begin by offering an overview of the core concepts and methods of TDA. Next, we survey the current applications of TDA in economics, emphasizing areas where it has proven useful such as the financial markets, risk management, and time series analysis. We also explore its limitations, particularly in terms of computational complexity and interpretability.

# 2: TDA Concepts and Methods