

Seminar in Data Science and AI

Professor Zhigang Yao

National University of Singapore, Singapore

Topic

Interaction of Statistics and Geometry: A New Landscape for Data Science

Date| Time

9 April 2025 (Wednesday) | 16:00 – 17:00 (HK Time)

Venue

Y305

Abstract:

While classical statistics primarily focuses on observations as real numbers or elements of real vector spaces, contemporary statistical challenges often involve more complex data types. These data are represented in spaces that, although not strictly Euclidean vector spaces, possess inherent geometric structures. The community exploring the interaction between statistics and geometry is expanding in both numbers and scope. The concept of manifold fitting traces back to H. Whitney's work in the early 1930s. The resolution of the Whitney extension problem has yielded new insights into data interpolation and inspired the formulation of the Geometric Whitney Problems. Specifically, given a set, we inquire: when can we construct a smooth d -dimensional submanifold to approximate the set, and how effectively can we estimate it in terms of distance and smoothness? In this talk, I will explore the manifold fitting problem, highlighting its modern insights and implications. Although various mathematical approaches have been proposed, many rely on restrictive assumptions, complicating the development of efficient and practical algorithms. As the manifold hypothesis-exploring non-Euclidean structures-remains a cornerstone of data science, further exploration of the manifold fitting problem is essential within the contemporary data science community. This discussion will be informed by recent work by Yao, Yau, and other co-authors, alongside ongoing research.

ALL ARE WELCOME