

Institut f. Geometrie

Gastvortrag

13.11.2019, 9:00 Uhr

Seminarraum 2, Kopernikusgasse 24/IV

Analysis of gene expression profiles via Wasserstein optimal transport on networks

CAROLINE MOOSMÜLLER

(Univ. California, San Diego)

Biological data sets, such as gene expressions or protein levels, are often high-dimensional and thus difficult to interpret. Finding important structural features and identifying clusters in an unbiased fashion is a core issue for understanding biological phenomena. In this talk, we describe an unsupervised data analysis methodology based on network analysis via Wasserstein optimal transport, dimension reduction with diffusion maps, and clustering with Mapper (TDA). Applied to gene expression profiles of the sarcomas in the Cancer Genome Atlas, we are able to recover the known subtypes. In addition, we find a new signature, mainly described by inactivation of the tumor suppressor genes p53 and p73, and discuss possible treatment based on its genetic profile.

This is joint work with J. C. Mathews, M. Pouryahya, I. G. Kevrekidis, J. O. Deasy, and A. Tannenbaum.

J. Wallner