Model-based microbiome data ordination: A variational approximation approach

Tao Wang

Shanghai Jiao Tong University, China
E-mail: neowangtao@sjtu.edu.cn

Abstract: The coevolution between human and bacteria colonizing the human body has profound implications for heath and development, with a growing body of evidence linking the altered microbiome composition with a wide array of disease states. Yet dimension reduction and visualization analysis of microbiome data is still in its infancy and many challenges exist. In this talk we introduce a general framework, Zero-Inflated Probabilistic PCA (ZIPPCA, pronounced zipcar), for dimension reduction and data ordination of multivariate abundance data, and propose an efficient variational approximation method for estimation, inference, and prediction. Extensive simulations show that the proposed method outperforms algorithm-based methods and compares favorably with existing model-based methods. We further apply our method to a gut microbiome dataset for visualization analysis of community composition across age and geography.