Mathematical Biology Seminar
Wednesday, March 3, 2021
Speaker: Audrey Fu, University of Idaho

Title: Imputation and causal network inference in genomics

Time: 9am
Zoom Link:
https://ucmerced.zoom.us/j/98050375649
Passcode: 172069

Abstract: Genomic data can be complex, large, noisy & sparse. Here I will discuss two problems we have worked on. The first deals with the highly sparse data from single-cell experiments of gene expression. These data contain a large number of zeros (>80%); many of these are missing values rather than no expression. Underlying these data are complex regulatory relationships among genes, as well as potentially many cell types with different gene expression profiles. We took a deep learning approach and designed imputation methods based on autoencoders. We generated synthetic data using real single-cell data to evaluate the performance, although the theoretical properties of autoencoders for imputation are yet to be understood.

The second problem deals with causal network inference: Can we learn about the complex regulatory relationships among genes directly from genomic data? Genetic variation makes this inference possible (under certain assumptions), as it provides randomization among the individuals -- this is known as the principle of Mendelian randomization in genetic epidemiology. We extended the interpretation of this principle to capture diverse causal relationships. We also developed an algorithm for learning causal networks based on the PC algorithm, a classical algorithm in computer science for inferring directed acyclic graphs.

Organized by:
Suzanne S. Sindi ssindi@ucmerced.edu &
Erica Rutter erutter2@ucmerced.edu

UC Merced's Mathematical Biology Seminar series is supported by an NSF-RTG grant (#1840265, PI: A. Kim) and associated with CRN: 16727.

Full Seminar Schedule: https://appliedmath.ucmerced.edu/node/52

Join our Mailing List: https://bit.ly/2LgjnT6
Join our Slack: https://bit.ly/3il4cIC