

Analysis and Applied Mathematics Seminar

Efficient distribution classification via optimal transport embeddings

Caroline Moosmueller (UCSD)

Abstract: Detecting differences and building classifiers between distributions, given only finite samples, are important tasks in a number of scientific fields. Optimal transport (OT) has evolved as the most natural concept to measure the distance between distributions, and has gained significant importance in machine learning in recent years.

There are some drawbacks to OT: Computing OT is usually slow, and it often fails to exploit reduced complexity in case the family of distributions is generated by simple group actions. In this talk, we discuss how optimal transport embeddings can be used to deal with these issues, both on a theoretical and a computational level.

In particular, we'll show how to embed the space of distributions into an L^2 -space via OT, and how linear techniques can be used to classify families of distributions generated by simple group actions in any dimension. The proposed framework significantly reduces both the computational effort and the required training data in supervised settings. We demonstrate the benefits in pattern recognition tasks in imaging and provide some medical applications.

Monday, October 25 at 4:00 PM in 636 SEO
--