

STATISTICS COLLOQUIUM

MONDAY, MARCH 24, 2014

TALK: 4:15 PM — SCIENCE CENTER RM. 705

RECEPTION: 3:50 PM — SCIENCE CENTER, 7TH FLOOR

“Nonparametric Inference for Unlabeled Graphs”

Peter Bickel
Department of Statistics
University of California, Berkeley

ABSTRACT

In a 2009 PNAS article, based on work of Aldous and Hoover (1981), B. and Chen proposed a general nonparametric framework, which could be related to the graphons of Lovasz et al. (2006), and large degree asymptotics for unlabeled random graphs. They studied fitting of block models viewable as “histogram approximations” in this framework. Recently, Airoldi, Costa and Chan (2013) and independently Olhede and Wolfe (2013) showed how to construct block model approximations to smooth densities of pairs of latent variables, conditional on existence of an edge, converging at appropriate speeds.

We discuss these results and argue that while they are hard to interpret statistically they imply statistically meaningful approximation. Moreover we argue that computationally efficient block model fitting procedures such as spectral and local methods can also be used for such approximation. Finally, we propose a principled cross validation method for choosing the “bandwidth” implicit in all such approximation procedures. We give some specific theory in this direction and an application to a proteomics network.

(Joint work with Patrick Wolfe and Sharmo Bhattacharyya.)